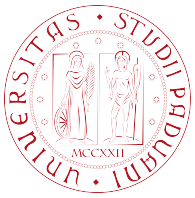


Università degli Studi di Padova

Sede Amministrativa: Università degli Studi di Padova
Dipartimento di Matematica

SCUOLA DI DOTTORATO DI RICERCA IN SCIENZE MATEMATICHE
INDIRIZZO: MATEMATICA COMPUTAZIONALE
CICLO XXVII

GAUSS QUADRATURE FOR LINEAR FUNCTIONALS AND NEW SEQUENCE TRANSFORMATIONS

Direttore della Scuola: Ch.mo Prof. Pier Paolo Soravia
Coordinatore d'indirizzo: Ch.ma Prof.ssa Michela Redivo Zaglia
Supervisore: Ch.ma Prof.ssa Michela Redivo Zaglia

Dottorando: Stefano Pozza

Riassunto

Questa tesi si compone di due parti. Nella prima parte viene presentata un'estensione della formula di quadratura di Gauss per l'approssimazione dei funzionali lineari quasi-definiti. Tale estensione viene costruita partendo dalla teoria dei polinomi ortogonali e in particolare dalla relazione tra le successioni di tali polinomi e alcune matrici dette matrici di Jacobi. La formula qui proposta, detta *quadratura di Gauss a n pesi*, soddisfa tutte le principali proprietà delle formula "classica" che definiremo *quadratura di Gauss a n nodi*. Inoltre, la tesi mostra come tale estensione possa essere calcolata tramite l'algoritmo di Lanczos non Hermitiano, al pari della formula a n nodi che può essere ottenuta tramite l'algoritmo di Lanczos Hermitiano. Al termine della prima parte sono presentati alcuni risultati preliminari relativi a una delle possibili applicazioni. Si tratta dell'approssimazione di *indici di centralità* di reti complesse, ovvero indici che stabiliscono quale nodo in un grafo è considerato più importante in termini di facilità di trasmissione di informazioni con altri nodi.

Nella seconda parte sono proposte alcune trasformazioni di successioni. Tali trasformazioni sono utilizzate al fine di ottenere, a partire da alcuni elementi di una successione data, un'altra successione che converge allo stesso limite ma a velocità maggiore. Infatti, spesso in analisi numerica e nella matematica applicata vi sono esempi di successioni, ottenute per esempio dai metodi iterativi, che convergono talmente lentamente da risultare inutili. È ben nota l'impossibilità di definire una trasformazione in grado di accelerare la convergenza di qualunque successione. Inoltre, usualmente le trasformazioni costruite per accelerare piccole classi di successione danno risultati migliori. Per questa ragione nel secondo capitolo di questa parte sono definite tre nuove trasformazioni in grado di accelerare una classe di successioni che estende quella relativa al noto processo di Aitken. Nella tesi vengono poi date condizioni necessarie affinché si abbia accelerazione della convergenza per la migliore delle tre trasformazioni proposte. Infine, tale trasformazione viene confrontata con altre trasformazioni. Da tale confronto si sono ottenuti risultati competitivi con alcuni dei più noti metodi di accelerazione (processo di Aitken, algoritmo ε, algoritmo θ, trasformazione di Levin).

Abstract

This thesis is divided into two parts. In the first one we present an extension of the Gauss quadrature formula for the approximation of the quasi-definite linear functionals. This extension is obtained using the orthogonal polynomials theory and, in particular, using the relation between sequences of these polynomials and some matrices called Jacobi matrices. We call this proposed formula *n-weight Gauss quadrature* and we show that it satisfies all the main properties of the "classical" formula, which we call *n-node Gauss quadrature*. Furthermore, we show that the proposed quadrature can be computed by the non-Hermitian Lanczos algorithm, in the same way in which the n-node Gauss quadrature can be computed by the Hermitian Lanczos algorithm. In the last chapter of the first part we present some preliminary results about possible applications. We approximate the *centrality indexes* of a complex network. These are indexes that measure the importance of a node in terms of communicability in a graph.

In the second part we propose some sequence transformations. Using sequence transformations we can use the elements of a sequence to obtain another sequence which converges faster to the same limit of the original one. Indeed, in numerical analysis and applied mathematics we often consider sequences arising, for example, from iterative methods that converge so slowly that they become useless. It has been proved that there cannot exist a transformation able to accelerate every sequence. Moreover, usually better results are given by transformations which are built to accelerate little classes of sequences. For this reason in the second chapter of this part we define three new transformations able to accelerate a class of sequences which extends the class of the well-known Aitken's process. We then consider the best of the three transformations and give some necessary conditions under which it accelerates the convergence of a given sequence. Finally, this transformation is compared with some of the most used transformations (Aitken's process, ε-algorithm, θ-algorithm, Levin type transformation) obtaining competitive results.

Acknowledgment

First of all, I would like to thank my advisor prof. Michela Redivo-Zaglia who encouraged me to follow my ideas (even when they brought me a bit far away from her present studies). I am grateful to her for providing me helpful advices, and for giving me the opportunities and the freedom to attend many meetings, conferences and meet important people for my formation. I would like to thank prof. Zdenek Strakoš. My visits in Prague were really important for the first part of this thesis and for me. I learned a lot from him, but most important, I always came back to Padova with a renewed enthusiasm for my work. Moreover, I would like to thank him for having introduced prof. Miroslav Pranić to me. Without prof. Pranić this thesis would not have been possible. The one week visit in Baja-Luka was extremely important and, most important, I would like to thank him for his friendly hospitality.

I am very grateful to prof. Gerard Meurant for introducing me to the studies of Gauss quadrature, moments and Lanczos algorithms. His helpful advices at the beginning and in the end of my studies were fundamental for this thesis and my work. For the same reason I would like to thank prof. Claude Brezinski for introducing me to the extrapolation methods studies. Moreover, thank to his lectures at the end of my master degree I began to consider to apply for a Ph.D in computational mathematics.

I am grateful to many friends among the Ph.D students I have met in these years in Padova and abroad, but I would like to especially thank Anna and Silvia who helped and supported me in many occasions.

Last but not least, I would like to express my gratitude to my family and my friends. Without the support and the help of my family (especially in this last year) it would have been impossible for me to finish my Ph.D studies. I hope that these hard times we went through will allow us to lay the basis for the future years. I would like to thank my brother Gianluca in particular, because we have been supporting each other again, as we will always do. Finally, old friends and new ones are really too many to stay on this page. However, I would like to thank everybody because everyday I walk with each of you through the several worlds I find on my way, and that is what has always given me the knowledge and the force I have needed. “La strada è l’unica salvezza”.

Contents

Introduction

Gauss quadrature and sequence transformations are two fundamental topics in numerical analysis. The first one is classical, however its extension to the approximation of linear functionals is not completely developed, even if many parts of this theory has been already stated for the case of quasi-definite linear functionals. In the first part of the thesis we present some results that try to improve this situation. The second topic considers a useful tool for the acceleration of slowly converging sequences: sequence transformations. In particular, in the last years many works were published on the effective use of sequence transformations; see [5, 24, 26, 30], and [19] in which many numerical techniques for the evaluation of power series expansions for special functions are investigated. Hence, it is of interest to introduce new sequence transformations.

This thesis extends and completes the results presented in the submitted paper [74] and in the published paper [14]. The first one concerns the approximation of a class of linear functionals through a Gauss quadrature-like rule that we investigate in Part I. The second one is about the acceleration of the convergence of sequences in a particular class with some sequence transformations (Part II).

Please, notice that the references are related to two different bibliographies, one for each part of the thesis.

Part I In the classical theory we are interested in the approximation of a Riemann, a weighted Riemann or a Riemann-Stieltjes integral

$$\int_{\mathbb{R}} f \mathrm{d}\mu,$$

with respect to a non-decreasing distribution function μ having finite limits at $\pm\infty$ and infinitely many points of increase (see [33], [54], [16], [17] and [84]). Since μ is of bounded variation, the integral exists for every continuous function f.

In the classical theory, we can introduce the well-known *Gauss Quadrature Rule*, which is the unique n-node quadrature with algebraic degree of exactness $2n-1$. The quadrature formula is usually obtained through orthogonal polynomials and their properties. A sequence of formal orthogonal polynomials $p_0, p_1, p_2, \ldots$ is a sequence such that $\int_{\mathbb{R}} p_i p \mathrm{d}\mu = 0$ for every polynomial p of degree lower than i. When we consider these kind of integrals, orthogonal polynomials are unique. In addition, every polynomial p_i has degree i and its roots are distinct. These last properties are fundamental for obtaining a Gauss quadrature for the considered integral.

When we consider an integral with respect to the measure μ, orthogonal polynomials and Gauss quadrature are strictly related with the set of real tridiagonal symmetric matrices with nonzero elements on their sub- and super-diagonal, which are usually known as *Jacobi matrices*. In particular, every finite sequence of orthogonal polynomials $p_0, \ldots, p_n$ can be associated with a Jacobi matrix J_n whose eigenvalues coincide with the roots of p_n. Moreover, these eigenvalues are the nodes of the n-node Gauss quadrature rule for the integral with respect to whom the polynomials are orthogonal. As remarked by Liesen and Strakoš in [63], Jacobi matrices are like a cornerstone between two wings of a building. The purpose of the first one is to approximate functions and integrals and it related to orthogonal polynomials, moments and continued fraction theories. The goal of the second one is matrix computations (solving linear systems, approximation of eigenvalues, ...) and it is related to vectors, vector spaces and matrices. Naturally, there exist many references about the classical theory of orthogonal polynomials and Gauss quadrature. In this thesis we will refer to [87, 15, 35, 36, 63].

In Part I our goal is to extend the Gauss quadrature to the approximation of a linear functional $\mathcal{L} : \mathcal{P} \to \mathbb{C}$, $\mathcal{P}$ being the space of complex polynomials. Indeed, it is well-known that we can define a sequence of orthogonal polynomials $p_0, \ldots, p_n$ with respect to $\mathcal{L}$ if the functional is quasi-definite. Unfortunately, these polynomials loose some of the important properties described before. For this reason extending the Gauss quadrature formula for a linear functional is not trivial and until now it has been done only under some restrictive assumptions (see [44, 78]). We achieve our goal under the assumption that $\mathcal{L}$ is quasi-definite and generalizing the expression of the quadrature rule. Furthermore, we show why we consider the introduced quadrature rule as the maximal extension of the Gauss quadrature concept for linear functionals. Moreover, we recover the link with Jacobi matrices extending the

definition of a Jacobi matrix in order to consider the complex case. This modify some properties of Jacobi matrices. Hence we need to recover or to give an extended version of the properties of real Jacobi matrices.

A discrete linear functional can be represented using a matrix $A \in \mathbb{C}^{n\times n}$ and two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ in the following way

$$\mathcal{L}(f) = \mathbf{u}^* f(A)\mathbf{v},$$

with f a matrix function. When A is Hermitian and $\mathbf{u} = \mathbf{v}$ it is classical to notice that the approximation of this bilinear form is equivalent to approximating an integral with respect to a nonnegative discrete measure. Hence, it can be approximated through Gauss quadrature. Let J_n be the Jacobi matrix obtained by the first n-th steps of the Hermitian Lanczos algorithm with input A and $\mathbf{v}$. It is well known that

$$\mathcal{L}(p) = \mathbf{e}_1^T p(J_n)\mathbf{e}_1,$$

for every polynomial p of degree $2n-1$. Moreover, $\mathbf{e}_1^T p(J_n)\mathbf{e}_1$ is equivalent to the n-node Gauss quadrature for the integral and J_n is the Jacobi matrix associated with a sequence of orthogonal polynomials with respect to $\mathcal{L}$.

Furthermore, the previous formula is valid even if A is a complex non-Hermitian matrix and $\mathbf{v} \neq \mathbf{u}$ are complex vectors. In this case, if $\mathcal{L}$ is quasi-definite, then J_n is complex and can be obtained by the non-Hermitian Lanczos algorithm; as proved, e.g, in [85]. Hence, in this thesis we investigate the relation between non-Hermitian Lanczos algorithm and the proposed extension of quadrature for linear functional.

Finally, we describe a numerical application in the complex networks field. Indeed, we will approximate the *subgraph centrality index* of a node in a graph approximating a bilinear form of the kind

$$\mathbf{e}_1^T \exp(A)\, \mathbf{e}_1,$$

with A real and non-symmetric, and $\exp(A)$ a matrix function.

We now summarize the contents of each chapter:

Chapter 1 We recall definitions and main properties of orthogonal polynomials with respect to a linear functional. In particular, we show some important properties of the positive definite case.

Chapter 2 Matrix functions can be defined in many different ways. We show the equivalence between the different definitions and prove useful properties.

Chapter 3 We give the definition of Jacobi matrix including the complex matrices. We prove some important spectral theorems and propositions, in particular regarding Jacobi matrices as tridiagonal matrices, and Jacobi matrices as symmetric matrices.

Chapter 4 Using the results of the previous chapters we introduce the *n-weight Gauss quadrature*, a quadrature rule for quasi-definite linear functionals with degree of exactness $2n - 1$.

Chapter 5 We show the relation between non-Hermitian Lanczos algorithm and the n-weight Gauss quadrature. In particular, we consider a formulation of the algorithm for the real non-symmetric case.

Chapter 6 We give an introduction to the complex networks theory in order to show the preliminary results of an application of the non-Hermitian Lanczos algorithm for the computation of the *subgraph centrality* of a node in a graph.

Part II In numerical analysis and applied mathematics we often consider sequences and series which can arise, for example, from iterative methods or other approximation technique. Frequently these sequences converge so slowly that they become useless. Hence, sequence transformations are fundamental since they potentially accelerate the convergence of a sequence. Indeed, we can try to transform a sequence to another one with better convergence properties.

Sequence transformation are based on extrapolation methods and have particular relations with the orthogonal polynomials theory and other topic, e.g., Padé approximation, continued fractions and projection methods. The literature on sequence transformation is ample; we refer to , e.g., [7, 11, 12, 15, 28, 32, 37]. It is well-known that there does not exist a transformation able to accelerate every sequence. Hence, it is of interest to introduce transformations specific for a class of sequences. Moreover, usually such tailored transformations produce better acceleration performance than more general ones.

The goal of the second part is to introduce three new sequence transformations whose kernel is

$$S_n = S + a_n \lambda^n,$$

with a_n a given sequence and S, λ unknowns. Clearly, it is a generalization of the Aitken' Δ^2 process kernel

$$S_n = S + a\lambda^n,$$

where S, a, λ are unknowns. To obtain these three transformations we follow a path similar to the one used by Brezinski and Redivo-Zaglia in [13]. In addition, we prove some acceleration properties for the transformation with the best performance. Then we compare this latter transformation with the ones proposed in [13] and with several important transformations, i.e, Aitken's Δ^2 process, Shanks' transformation, θ-algorithm and Levin type transformation. Furthermore, we report the results obtained by the acceleration of the sequence of the approximations of the digamma function using our proposed transformation. Finally, we observe what happens when the acceleration is not guaranteed by the proved theorems, i.e., the cases of logarithmically convergent or monotone sequences. The numerical experiments show that the proposed transformation can be competitive for sequences near its kernel with respect to all the other considered transformations. Moreover, we have some particularly interesting results in the case of diverging sequences, since the proposed transformation seems to have a good semiconvergent behavior.

We now summarize the contents of each chapter:

Chapter 1 We introduce the basic theory of sequence transformations and extrapolation methods. Moreover, we define and recall some properties of Aitken's Δ^2 process, Shanks' transformation, θ-algorithm and Levin type transformation.

Chapter 2 We define three new sequence transformations and analyze their convergence and acceleration properties. Then, in several numerical examples we compare the best of these new transformations with the ones introduced in the previous chapter. In particular, we test it in the acceleration of a sequence approximating the digamma function.

Part I

Gauss Quadrature for Quasi-definite Linear Functionals

CHAPTER 1

Linear Functionals and Orthogonal Polynomials

The history of orthogonal polynomials began in the nineteenth century from investigations on a particular kind of continued fractions named after Stieltjes. The theory is classical, and there are many books on this topic. As a basic reference we consider, besides the classical monograph by Szegö [87], the beautiful summary by Chihara [15]. We also mention, for the computational aspects, the book by Gautschi [35].

Let $\mathcal{L}$ be a *linear* functional on the space of (complex) polynomials, $\mathcal{L} : \mathcal{P} \to \mathbb{C}$. We say that $\pi_0, \pi_1, \ldots, \pi_k$ is a sequence of *formal orthogonal polynomials* if

$$\pi_j \in \mathcal{P}_j \quad \text{and} \quad \mathcal{L}(p\,\pi_j) = 0, \quad \forall p \in \mathcal{P}_{j-1}, \quad \text{for } j = 1, 2, \ldots, k,$$

where $\mathcal{P}_j$ is the space of polynomials of degree at most j.

In the classical case (see [33], [54], [16], [17] and [84]) $\mathcal{L}$ is the Riemann, the weighted Riemann or the more general Riemann-Stieltjes integral with respect to a non-decreasing distribution function μ defined on the real axis having finite limits at $\pm\infty$ and infinitely many points of increase. Since μ is of bounded variation, the integral $\int f \mathrm{d}\mu$ exists for every continuous function f. Moreover, under these assumptions, and with $\mathcal{L}(f) = \int f \mathrm{d}\mu$, formal orthogonal polynomials π_j have some additional properties: they exist, they are unique up to a nonzero constant factor, they satisfy a three-term recurrence relation, π_j is of degree j and $\mathcal{L}(\pi_j^2) \neq 0$ for $j = 0, 1, \ldots$ These properties can be extended to sequences of orthogonal polynomials with respect to a partic-

ular class of linear functionals (quasi-definite linear functionals). In Section 1.1 we present the main results about this kind of orthogonal polynomials sequences. In particular, we focus on their zeros, on the fact that they satisfy a three-term recurrence relation and on their relation with *Jacobi matrices*. In Section 1.2 we discuss the case of positive-definite linear functionals and their properties.

1.1 Orthogonal Polynomials

The term *orthogonal polynomials* usually refers to polynomials orthogonal with respect to an inner product

$$\langle p, q\rangle = \int_{\mathbb{R}} pq\, d\mu,$$

with μ a positive Borel measure. We consider a more general case in which the polynomials are orthogonal with respect to a linear functional; see [15]. Let the linear functional $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ defined on the space of (complex) polynomials $\mathcal{P}$. The functional $\mathcal{L}$ is fully determined by its values on monomials, called moments,

$$\mathcal{L}(x^k) = m_k, \quad k = 0, 1, \ldots . \tag{1.1}$$

From now on $\mathcal{L}$ will always refer to this kind of linear functionals.

Definition 1.1. *A sequence of polynomials $\pi_0, \pi_1, \ldots, \pi_k$ satisfying the conditions*

1. *$deg(\pi_j) = j$ (π_j is of degree j),*
2. *$\mathcal{L}(\pi_i\, \pi_j) = 0$, $i < j$,*
3. *$\mathcal{L}(\pi_j^2) \neq 0$,*

is called a sequence of orthogonal polynomials with respect to the linear functional $\mathcal{L}$.

The linearity of $\mathcal{L}$ and the condition 3. implies $\pi_0(x) \neq 0$ and $m_0 \neq 0$. Furthermore, the conditions 2.–3. are equivalent respectively to the following:

$$\mathcal{L}(p\pi_j) = 0,\ \forall p \in \mathcal{P}_{j-1}, \qquad \text{and} \qquad \mathcal{L}(p\pi_j) \neq 0,\ \text{if } \deg(p) = j. \tag{1.2}$$

Indeed, since $\pi_0, \ldots, \pi_{j-1}$ is a basis of $\mathcal{P}_{j-1}$ p can be written as $p = \sum_{k=0}^{j-1} a_k \pi_k$. Hence, if $\mathcal{L}(\pi_i\, \pi_j) = 0$ then $\mathcal{L}(p\pi_j) = 0$, for all $p \in \mathcal{P}_{j-1}$.

Moreover, if p has degree j and $\mathcal{L}(\pi_j^2) \neq 0$ then the same argument gives $\mathcal{L}(p\pi_j) \neq 0$. The converse implications are trivial.

As we mentioned above, $\pi_0, \pi_1, \ldots, \pi_n$ are a basis for $\mathcal{P}_n$. Hence, given a polynomial $p \in \mathcal{P}_n$ we can rewrite it as

$$p = \sum_{k=0}^{n} a_k \pi_k,$$

with

$$a_k = \frac{\mathcal{L}(p\pi_k)}{\mathcal{L}(\pi_k^2)} \quad \text{for } k = 0, \ldots, n. \tag{1.3}$$

The expression for a_k follows since for $k = 0, \ldots, n$ $\mathcal{L}(\pi_k^2) \neq 0$ and

$$\mathcal{L}(p\pi_k) = \sum_{j=0}^{n} a_k \mathcal{L}(\pi_j \pi_k) = a_k \mathcal{L}(\pi_k^2).$$

Providing that they exist, $\pi_n(x)$, $n = 1, 2, \ldots, k$ are uniquely determined up to a nonzero constant factor. Indeed, if both π_n and $\tilde{\pi}_n$ are orthogonal polynomials then $\tilde{\pi}_n = \sum_{k=0}^{n} a_k \pi_k$, and by (1.3) we conclude $\tilde{\pi}_n = a_n \pi_n$.

About the question of existence of orthogonal polynomials, this is considered, for example, in [15, Chapter I]; for the case of classical orthogonal polynomials see also Theorem 2.1.1 and pages 24 and 25 of [87]. In the following we will discuss this issue using the Hankel determinants Δ_j of the matrices of moments (see (1.1)),

$$\Delta_j = \begin{vmatrix} m_0 & m_1 & \ldots & m_j \\ m_1 & m_2 & \ldots & m_{j+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_j & m_{j+1} & \ldots & m_{2j} \end{vmatrix}. \tag{1.4}$$

Definition 1.2. *A linear functional $\mathcal{L}$ for which the first $k+1$ Hankel determinants are nonzero, i.e. $\Delta_j \neq 0$ for $j = 0, 1, \ldots, k$, is called quasi-definite on the space $\mathcal{P}_k$ of polynomials of degree at most k.*

Theorem 1.3. *[15, Chapter I, Theorem 3.1] A sequence $\pi_0, \ldots, \pi_k$ of orthogonal polynomials with respect to $\mathcal{L}$ exists if and only if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_k$.*

Proof. Assume that $\pi_0, \ldots, \pi_{n-1}$ exist and write

$$\pi_n(x) = \sum_{i=0}^{n} a_{n,i} x^i.$$

orthogonality conditions (1.2) give $\mathcal{L}(x^n\pi_n) = b_n \neq 0$ and

$$\mathcal{L}(x^k\pi_n) = \sum_{i=0}^{n} a_{n,i}m_{i+k} = \begin{cases} 0, & \text{for } k = 0, \dots, n-1 \\ b_n, & \text{for } k = n; \end{cases}$$

. These conditions are equivalent to the linear systems

$$\begin{bmatrix} m_0 & m_1 & \cdots & m_n \\ m_1 & m_2 & \cdots & m_{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ m_n & m_{n+1} & \cdots & m_{2n} \end{bmatrix} \begin{bmatrix} a_{n,0} \\ a_{n,1} \\ \vdots \\ a_{n,n} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ b_n \end{bmatrix} \tag{1.5}$$

If the n-th orthogonal polynomial exists, then it is uniquely determined by b_n. Then the system (1.5) has a unique solution, thus $\Delta_n \neq 0$. Conversely, let $\Delta_n \neq 0$. If we choose a value $b_n \neq 0$, then the system (1.5) has a unique solution. Moreover, for $n = 0, \dots, k$,

$$a_{n,n} = \frac{b_n\Delta_{n-1}}{\Delta_n} \neq 0, \tag{1.6}$$

with $\Delta_{-1} = 1$. Hence, p_n has degree n. □

A fundamental property of orthogonal polynomials is the simple relation that holds between any three consecutive monic polynomials of the sequence; see [87, Theorem 3.2.1] and [15, Theorem 4.1]. From now on π is always used for monic orthogonal polynomials.

Theorem 1.4. *Let $\mathcal{L}$ be a quasi-definite linear functional on $\mathcal{P}_n$ and $\pi_0, \dots, \pi_k$ the monic orthogonal polynomials with respect to $\mathcal{L}$, then*

$$\pi_n(x) = (x - \delta_{n-1})\pi_{n-1} - \eta_{n-1}\pi_{n-2}, \quad n = 1, 2, \dots, k \tag{1.7}$$

where we set $\eta_0 = m_0$, while the other elements are defined as

$$\delta_{n-1} = \frac{\mathcal{L}(x\pi_{n-1}^2)}{\mathcal{L}(\pi_{n-1}^2)}, \quad \eta_{n-1} = \frac{\mathcal{L}(\pi_{n-1}^2)}{\mathcal{L}(\pi_{n-2}^2)} \neq 0, \quad \pi_{-1} \equiv 0,\ \pi_0 \equiv 1;$$

Proof. $x\pi_{n-1}(x)$ is a polynomial of degree n, thus it can be written

$$x\pi_{n-1}(x) = \sum_{i=0}^{n} a_{n-1,i}\pi_i,$$

with

$$a_{n-1,i} = \frac{\mathcal{L}(x\pi_{n-1}(x)\pi_i(x))}{\mathcal{L}(\pi_i^2(x))} \quad \text{for } i = 0, \dots, n.$$

Since $x\pi_i(x)$ has degree $i+1$, $a_{n-1,i} = 0$ for $i = 0, \ldots, n-3$. Moreover, $x\pi_{n-1}(x)$ is monic, thus $a_{n-1,n} = 1$. Then we have

$$x\pi_{n-1}(x) = \pi_n + a_{n-1,n-1}\pi_{n-1} + a_{n-1,n-2}\pi_{n-2}.$$

Setting $\delta_{n-1} = -a_{n-1,n-1}$, $\eta_{n-1} = -a_{n-1,n-2}$, $\eta_0 = m_0$, $\pi_{-1}(x) = 0$, and $\pi_0(x) = 1$ we obtain (1.7).

Multiplying (1.7) by π_{n-1} and applying $\mathcal{L}$ we obtain

$$0 = \mathcal{L}(x\pi_{n-1}^2(x)) - \delta_{n-1}\mathcal{L}(\pi_{n-1}^2),$$

from which we easily obtain δ_{n-1}. Multiplying (1.7) by π_{n-2} and using a similar argument we conclude the proof. □

Unlike in the classical case, in which the functional is an integral, for $\mathcal{L}$ quasi-definite the coefficients of the associated orthogonal polynomials are not necessarily real, the coefficients in the three-term recurrence relation are, in general, complex, and zeros of the orthogonal polynomials can be complex and multiple.

We obtain a sequence of *orthonormal polynomials* $\tilde{p}_j$ with the normalization of the individual polynomials π_j. The normalized sequences are unique up to multiplication by (-1), and one particular sequence within the whole family can be expressed as

$$\tilde{p}_j(x) = \frac{\pi_j(x)}{\sqrt{\mathcal{L}(\pi_j^2)}} = \frac{\pi_j(x)}{\sqrt{\eta_0\eta_1\cdots\eta_j}},\ j = 1, 2, \ldots, k\,, \tag{1.8}$$

where we take $arg(\sqrt{c}) \in (-\pi/2, \pi/2]$, i.e., consider the principal value of the square root. Hence, if there exist a sequence of monic orthogonal polynomials $\pi_0, ..., \pi_k$, then there are 2^{k+1} associated sequences of orthonormal polynomials. These sequences clearly differ in the computation of the complex square roots of the individual coefficients $\eta_1, \ldots, \eta_k$. The three-term recurrence relation for orthonormal polynomials $\tilde{p}_0, \ldots, \tilde{p}_n$, $n \leq k$, can be written as

$$x\begin{bmatrix} \tilde{p}_0(x) \\ \tilde{p}_1(x) \\ \vdots \\ \tilde{p}_{n-1}(x) \end{bmatrix} = J_n \begin{bmatrix} \tilde{p}_0(x) \\ \tilde{p}_1(x) \\ \vdots \\ \tilde{p}_{n-1}(x) \end{bmatrix} + \sqrt{\eta_n}\begin{bmatrix} 0 \\ 0 \\ \vdots \\ \tilde{p}_n(x) \end{bmatrix}, \tag{1.9}$$

where J_n is the (complex) tridiagonal symmetric matrix

$$J_n = \begin{bmatrix} \delta_0 & \sqrt{\eta_1} & & & \\ \sqrt{\eta_1} & \delta_1 & \sqrt{\eta_2} & & \\ & \sqrt{\eta_2} & \delta_2 & \ddots & \\ & & \ddots & \ddots & \sqrt{\eta_{n-1}} \\ & & & \sqrt{\eta_{n-1}} & \delta_{n-1} \end{bmatrix}. \tag{1.10}$$

From (1.9) we see that the zeros λ_i, $i = 1, \ldots, n$, of $\tilde{p}_n$ are the eigenvalues of J_n, with

$$\mathbf{w_i} = [\tilde{p}_0(\lambda_i),\ \tilde{p}_1(\lambda_i),\ \ldots,\ \tilde{p}_{n-1}(\lambda_i)]^T,\ i = 1, \ldots, n, \tag{1.11}$$

the associated eigenvectors.

Theorem 1.4 can be extended to the general case of a sequence of *orthogonal polynomials* $p_0, p_1, \ldots$. Indeed, it satisfies the three-term recurrence relationship of the form

$$\beta_n p_n(x) = (x - \alpha_{n-1}) p_{n-1}(x) - \gamma_{n-1} p_{n-2}(x), \quad \text{for} \quad n = 1, 2, \ldots, \tag{1.12}$$

where we set $\gamma_0 = 0$, $p_{-1} \equiv 0$, $p_0 \equiv c$ (c is a given complex number different from zero) and

$$\alpha_{n-1} = \frac{\mathcal{L}(x p_{n-1}^2)}{\mathcal{L}(p_{n-1}^2)},\ \beta_n = \frac{\mathcal{L}(x p_{n-1} p_n)}{\mathcal{L}(p_n^2)},\ \gamma_{n-1} = \frac{\mathcal{L}(x p_{n-2} p_{n-1})}{\mathcal{L}(p_{n-2}^2)}, \tag{1.13}$$

(see [87, Theorem 3.2.1], [15, p. 19], [6, Theorem 2.4]). Providing that $p_0, p_1, \ldots, p_n$ exist, all coefficients β_j and γ_{j-1}, for $j = 0, \ldots, n$, are different from zero by 3. of Definition 1.1. The recurrences (1.12) can be written in matrix form

$$x \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{n-1}(x) \end{bmatrix} = T_n \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{n-1}(x) \end{bmatrix} + \beta_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ p_n(x) \end{bmatrix}. \tag{1.14}$$

Now, T_n is a tridiagonal complex matrix

$$T_n = \begin{bmatrix} \alpha_0 & \beta_1 & & \\ \gamma_1 & \alpha_1 & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & \gamma_{n-1} & \alpha_{n-1} \end{bmatrix}.$$

On the other hand, we can obtain orthogonal polynomials with respect to a linear functional from a three-term recurrence relation defining a sequence of polynomials. This was shown firstly for the classical case by Favard in [27] and for the general case, for example, in [15, Chapter I, Theorem 4.4]; see also the survey [64, Theorem 2.14]. Here we adapt the proof so that we can consider three-term recurrence relations with a finite number of polynomials $p_0, \ldots, p_{k+1}$.

Theorem 1.5 (Favard)**.** *Let $p_0, \ldots, p_{k+1}$ polynomials satisfying*

$$b_{n+1}p_{n+1}(x) = (x - a_n)p_n(x) - c_n p_{n-1}(x),\ n = 0, 1, \ldots, k \tag{1.15}$$

where

$$p_{-1} \equiv 0,\ p_0 \equiv c,\ c_0 = 0,\ a_n, b_n, c_n, d \in \mathbb{C},\ b_n, c_n, c \neq 0,$$

then there exists a linear functional $\mathcal{L} : \mathcal{P}_{2k+1} \to \mathbb{C}$ quasi-definite on $\mathcal{P}_k$ such that $p_0, p_1, \ldots, p_k$ are orthogonal polynomials with respect to $\mathcal{L}$.

In other words, providing that $c, b_n, c_n \neq 0$, polynomials generated by (1.15) are always orthogonal polynomials. In addition, they are orthonormal if and only if $c_n = b_n$ and p_0 is such that $\mathcal{L}(p_0^2) = 1$.

Proof. Since $b_n \neq 0$ for $n = 1, \ldots, k$ and $p_0 = c$, polynomial p_n has degree n for $n = 0, \ldots, k+1$. Let $\mathcal{L} : \mathcal{P}_{2k} \to \mathbb{C}$ a linear functional defined by the conditions

$$\mathcal{L}(p_0) = 1, \tag{1.16}$$

$$\mathcal{L}(p_n) = 0 \quad \text{for } n = 1, \ldots, k, \tag{1.17}$$

$$\mathcal{L}(x^j p_{k+1}) = 0 \quad \text{for } j = 0, \ldots, k. \tag{1.18}$$

The polynomial $x^j p_{k+1}$ has degree $k + 1 + j$ for $j = 0, \ldots, k-1$, thus $p_0, \ldots, p_k, p_{k+1}, xp_{k+1}, \ldots, x^k p_{k+1}$ is a basis for $\mathcal{P}_{2k+1}$. This means that $\mathcal{L}$ is well defined by the previous conditions.

We can rewrite (1.15) as

$$xp_n(x) = -b_{n+1}p_{n+1}(x) + a_n p_n(x) + c_n p_{n-1}(x),\ n = 0, 1, \ldots, k, \tag{1.19}$$

which combined with assumption (1.17) gives

$$\mathcal{L}(xp_n(x)) = 0,\ n = 2, \ldots, k-1.$$

Assumptions (1.18) extend this equation for $n = k$. Multiplying (1.19) by x and using the previous results gives

$$\mathcal{L}(x^2 p_n(x)) = 0,\ n = 3, \ldots, k.$$

Repeating this argument we obtain

$$\mathcal{L}(x^j p_n(x)) = 0, \quad \text{for } 0 \leq j < n \leq k.$$

Multiplying again (1.19) by x^{n-1}, using the orthogonality property just obtained and using (1.18) we get

$$\mathcal{L}(x^n p_n) = c_n \mathcal{L}(x^{n-1} p_{n-1}), \quad \text{for } n = 1, \ldots, k,$$

which, with assumption (1.16), gives

$$\mathcal{L}(x^n p_n) = c_1 c_2 \cdots c_n, \quad \text{for } n = 1, \ldots, k-1.$$

Since $c_0, \ldots, c_n \neq 0$ this conclude the proof. □

Moreover, Theorem 1.5 says that for any tridiagonal matrix T_n without any zero components on the sub- and super-diagonal there exists a linear functional $\mathcal{L}$ quasi-definite on $\mathcal{P}_{n-1}$ for which T_n is determined by the moments $m_0, \ldots, m_{2n-1}$ of $\mathcal{L}$. We clarify the non uniqueness of determining T_n from the moments of $\mathcal{L}$ with the following statement.

Proposition 1.6. *Let T_n and $\widehat{T}_n$ be two tridiagonal matrices without zero components on the sub- and super-diagonal. Then, T_n and $\widehat{T}_n$ are determined by the first $2n$ moments of the same linear functional which is quasi-definite on $\mathcal{P}_{n-1}$ if and only if T_n and $\widehat{T}_n$ are similar matrices such that $T_n = D^{-1}\widehat{T}_n D$, where D is an invertible diagonal matrix.*

Proof. The proof uses formula (1.14) and the observation that two sets of polynomials $p_0, \ldots, p_{n-1}$ and $\hat{p}_0, \ldots, \hat{p}_{n-1}$ are orthogonal with respect to the same linear functional if and only if

$$\begin{bmatrix} \hat{p}_0(x) \\ \vdots \\ \hat{p}_{n-1}(x) \end{bmatrix} = D \begin{bmatrix} p_0(x) \\ \vdots \\ p_{n-1}(x) \end{bmatrix},$$

where D is an invertible diagonal matrix.

We first assume that T_n and $\widehat{T}_n$ are two matrices determined by the same moments of the linear functional $\mathcal{L}$ quasi-definite on $\mathcal{P}_{n-1}$. The matrices T_n and $\widehat{T}_n$ determine two sequences of orthogonal polynomials that we name respectively $p_0, \ldots, p_{n-1}$ and $\hat{p}_0, \ldots, \hat{p}_{n-1}$. Using the recurrence relation (1.12) we can define the polynomial

$$q_n = (x - \alpha_{n-1}) p_{n-1} - \gamma_{n-1} p_{n-2}. \tag{1.20}$$

and analogously the polynomial $\hat{q}_n$. The recurrence relation (1.14) for the polynomials $\hat{p}_0, \ldots, \hat{p}_{n-1}$ and $\hat{q}_n$ then gives

$$xD \begin{bmatrix} p_0(x) \\ \vdots \\ p_{n-1}(x) \end{bmatrix} = \widehat{T}_n D \begin{bmatrix} p_0(x) \\ \vdots \\ p_{n-1}(x) \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ \hat{q}_n(x) \end{bmatrix}. \tag{1.21}$$

Hence, we obtain that $T_n = D^{-1}\widehat{T}_n D$ and $q_n = \hat{q}_n/d_n$, with d_n the last diagonal element of D.

Vice versa, putting $T_n = D^{-1}\widehat{T}_n D$ in (1.14) and multiplying from the left by D we get (1.21) which means that we obtain two sequences of orthogonal polynomials such that $[\hat{p}_0, \ldots, \hat{p}_{n-1}]^T = D[p_0, ..., p_{n-1}]^T$. □

Remark 1.7. *Using (1.20) and (1.14) we can show that q_n has degree n and its zeros are the eigenvalues of T_n (analogously to the positive definite linear functional case in [63, Sections 3.2.1 and 3.4.1]). Moreover, q_n is orthogonal to $\mathcal{P}_{n-1}$.*

In the following the elements of $\widehat{T}_n$ are marked with a hat.

Corollary 1.8. *Let T_n and $\widehat{T}_n$ be two tridiagonal matrices without zero components on the sub- and super-diagonals. T_n and $\widehat{T}_n$ are determined by the first $2n$ moments of the same linear functional if and only if*

- $\alpha_i = \hat{\alpha}_i$ *for* $i = 0, \ldots, n-1$;
- $\beta_i\gamma_i = \hat{\beta}_i\hat{\gamma}_i$ *for* $i = 1, \ldots, n-1$.

Proof. By Proposition 1.6 we know that T_n and $\widehat{T}_n$ are determined by the first $2n$ moments of the same linear functional if and only if $T_n = D^{-1}\widehat{T}_n D$, with $D = \operatorname{diag}(d_1, \ldots, d_n)$ an invertible diagonal matrix. We first assume that $T_n = D^{-1}\widehat{T}_n D$. Comparing the corresponding entries of matrices T_n and $D^{-1}\widehat{T}_n D$ we get $\alpha_i = \hat{\alpha}_i$, for $i = 0, \ldots, n-1$, as well as $\gamma_i = (d_i/d_{i+1})\hat{\gamma}_i$ and $\beta_i = (d_{i+1}/d_i)\hat{\beta}_i$ for $i = 1, \ldots, n-1$. Thus we see that $\gamma_i\beta_i = \hat{\gamma}_i\hat{\beta}_i$ for $i = 1, \ldots, n-1$.

Vice versa, if $\alpha_i = \hat{\alpha}_i$ for $i = 0, \ldots, n-1$ and $\beta_i\gamma_i = \hat{\beta}_i\hat{\gamma}_i$ for $i = 1, \ldots, n-1$, then the diagonal matrix $D = \operatorname{diag}(d_1, \ldots, d_n)$ whose elements are $d_1 = 1$ and

$$d_i = \frac{\beta_1\beta_2\cdots\beta_{i-1}}{\hat{\beta}_1\hat{\beta}_2\cdots\hat{\beta}_{i-1}} = \frac{\hat{\gamma}_1\hat{\gamma}_2\cdots\hat{\gamma}_{i-1}}{\gamma_1\gamma_2\cdots\gamma_{i-1}}, \quad \text{for } i = 2, \ldots, n,$$

gives $T_n = D^{-1}\widehat{T}_n D$. □

Moreover, if T_n is a tridiagonal matrix with nonzero entries on sub- and super-diagonal, then it is similar to a *complex tridiagonal symmetric matrix* J_n. Indeed, the diagonal matrix $D = \operatorname{diag}(d_1, \ldots, d_n)$ with

$$d_1 = 1, \quad d_i = \left(\frac{\gamma_1 \gamma_2 \ldots \gamma_{i-1}}{\beta_1 \beta_2 \ldots \beta_{i-1}} \right)^{1/2}, \quad \text{for } i = 2, \ldots, n, \tag{1.22}$$

gives the similarity transformation we need. This result is well-known in the classical case for which J_n is a real tridiagonal symmetric matrix ([95, pp. 335-336]). We want to stress out that in the case of quasi-definite linear functionals the matrix J_n is, in general, complex. Hence J_n is symmetric but it may not be a Hermitian matrix.

Corollary 1.8 implies that there exist 2^{n-1} different tridiagonal symmetric matrices J_n determined by the moments $m_0, \ldots, m_{2n-1}$. In fact, two symmetric tridiagonal matrices J_n and $\widehat{J}_n$ with nonzero entries on the sub-diagonal (or super-diagonal) are determined by the first $2n$ moments of a linear functional if and only if they have the same diagonal and $\beta_i = \pm\hat{\beta}_i$ for $i = 1, \ldots, n-1$. Notice that this correspond with the nonuniqueness of the sequences of orthonormal polynomials mentioned above in this section.

We recall the following important identities for orthogonal polynomials; see [15, Chapter I, Theorem 4-5], [87, Theorem 3.2.2].

Theorem 1.9 (Christoffel-Darboux identity)**.** *Let $\pi_0, \ldots, \pi_k$ be monic orthogonal polynomials, then for $n = 0, \ldots, k-1$*

$$\sum_{j=0}^{n} \frac{\pi_j(x)\pi_j(y)}{K_j} = \frac{1}{K_n} \frac{\pi_{n+1}(x)\pi_n(y) - \pi_{n+1}(y)\pi_n(x)}{x - y}, \tag{1.23}$$

with $K_j = \mathcal{L}(\pi_j^2) = \eta_0 \eta_1 \ldots \eta_j, j = 0, \ldots, n$.

Proof. The three-term recurrence relation (1.7) gives

$$\begin{aligned} x\pi_i(x)\pi_i(y) &= \pi_{i+1}(x)\pi_i(y) + \delta_i \pi_i(x)\pi_i(y) + \eta_i \pi_{i-1}(x)\pi_i(y) \\ y\pi_i(y)\pi_i(x) &= \pi_{i+1}(y)\pi_i(x) + \delta_i \pi_i(y)\pi_i(x) + \eta_i \pi_{i-1}(y)\pi_i(x), \end{aligned}$$

for $i = 0, \ldots, k-1$. Subtracting the second equation from the first produces

$$\begin{aligned} (x - y)\pi_i(x)\pi_i(y) = \pi_{i+1}(x)\pi_i(y) - \pi_{i+1}(y)\pi_i(x) \\ - \eta_i(\pi_{i-1}(y)\pi_i(x) - \pi_{i-1}(x)\pi_i(y)) \end{aligned}$$

Let $F_i(x, y)$ be the right-hand side of equation (1.23). The previous equality becomes

$$\frac{\pi_i(x)\pi_i(y)}{K_i} = F_i(x, y) - F_{i-1}(x, y), \quad \text{for } i = 0, \ldots, k-1.$$

Summing the last equation for $j = 0, \ldots, n$ finishes the proof. □

Finally, rewriting the numerator of the right-hand side of (1.23) as

$$(\pi_{n+1}(x) - \pi_{n+1}(y))\pi_n(x) - (\pi_n(x) - \pi_n(y))\pi_{n+1}(x),$$

and letting $y \to x$ gives

$$\sum_{j=0}^{n} \frac{\pi_j(x)^2}{K_j} = \frac{\pi'_{n+1}(x)\pi_n(x) - \pi'_n(x)\pi_{n+1}(x)}{K_n}, \tag{1.24}$$

for $n = 0, \ldots, k-1$.

1.2 Positive Definite Linear Functionals

In this section we present some well-known facts about the positive definite case. Let $\mathcal{L}$ be a linear functional, $m_0, m_1, m_2, \ldots$ its moments and $\Delta_0, \Delta_1, \Delta_2, \ldots$ its Hankel determinants (1.4), In the same spirit of Definition 1.2 we define positive definite functionals (see for example [15, Chapter I, Definition 3.1]).

Definition 1.10. *The linear functional $\mathcal{L}$ is said to be positive definite on $\mathcal{P}_k$ if $m_s \in \mathbb{R}$ for $s = 0, \ldots, 2k$ and $\Delta_j > 0$ for $j = 0, \ldots, k$.*

By Theorem 1.3 given a linear functional positive definite on $\mathcal{P}_k$ there exist $p_0, \ldots, p_k$ orthogonal polynomials with respect to $\mathcal{P}_k$. Moreover, (1.5) implies that the coefficient of the polynomials $p_0, \ldots, p_k$ are real.

We are going to show that, as it is well-known, the classical theory of orthogonal polynomials is equivalent to the positive definite case. We say that a real polynomial f (i.e. a polynomial with real coefficients) is non-negative if $f(x) \geq 0$ for all real variable x. We recall the following lemma; see [15, p. 15].

Lemma 1.11. *Let f be a non-negative polynomial of degree n. Then there exist real polynomials p and q such that*

$$f(x) = p^2(x) + q^2(x),$$

with p^2 and q^2 of degree at most n.

Proof. If f is non-negative and has real coefficients, then its real zeros have even multiplicity and the non-real zeros occur in conjugate pairs. Then we can rewrite f as

$$f(x) = r^2(x) \prod_{k=1}^{n} (x - a_k - ib_k)(x - a_k + ib_k),$$

with r a real polynomial (with the same zeros and multiplicities as the real zeros of f), a_k, b_k real numbers and i the imaginary unit. The products can be rewritten as $(s(x) + it(x))(s(x) - it(x))$ with s, t real polynomials given by

$$\prod_{k=1}^{n}(x - a_k - ib_k) = s(x) + it(x).$$

The proof finishes since

$$f(x) = r^2(x)(s^2(x) + t^2(x)).$$

Moreover, clearly r^2s^2 and r^2t^2 have degree at most n. □

Theorem 1.12. *The linear functional $\mathcal{L}$ is positive definite on $\mathcal{P}_k$ if and only if $\mathcal{L}(f) > 0$ for every nonzero and nonnegative real polynomial f from $\mathcal{P}_{2k}$.*

Proof. Let $\mathcal{L}$ be positive definite on $\mathcal{P}_k$. Then, there exist $\pi_0, \ldots, \pi_k$ monic orthogonal polynomials with respect to $\mathcal{L}$. Above in this section we have noticed that the monic polynomials $\pi_0, \ldots, \pi_k$ have real coefficients, moreover they are a basis for $\mathcal{P}_k$. Thus, if p is a real polynomial of degree $n \leq k$ then

$$p = \sum_{j=0}^{n} a_j \pi_j,$$

with $a_0, \ldots, a_n$ real coefficients and $a_n \neq 0$. By (1.6) and since π_n is monic

$$b_j = \frac{\Delta_j}{\Delta_{j-1}}, \quad \text{for } j = 0, \ldots, k,$$

with $\Delta_{-1} = 1$ and $b_j = \mathcal{L}(x^j \pi_j)$. Hence

$$\mathcal{L}(\pi_j^2) = \mathcal{L}(x^j \pi_j) = \frac{\Delta_j}{\Delta_{j-1}} > 0, \quad \text{for } j = 0, \ldots, k. \tag{1.25}$$

Therefore

$$\mathcal{L}(p^2) = \sum_{i,j=0}^{n} a_i a_j \mathcal{L}(\pi_i \pi_j) = \sum_{j=0}^{n} a_j^2 \mathcal{L}(\pi_j^2) > 0.$$

Lemma 1.11 implies $\mathcal{L}(f) > 0$ for every nonzero and nonnegative real polynomial f from $\mathcal{P}_{2k}$.

Conversely, for $n = 0, \ldots, k$ $m_{2n} = \mathcal{L}(x^{2n}) > 0$ and since

$$0 < \mathcal{L}[(x+1)^{2n}] = \sum_{j=0}^{2n} \binom{2n}{j} m_{2n-j}$$

m_{2n-1} is real by induction. Using the three-term recurrence relation (1.7) it is easy to see that we can build a sequence of real monic orthogonal polynomials $\pi_0, \ldots, \pi_k$ with respect to $\mathcal{L}$. Indeed, δ_{n-1} is real and $\eta_{n-1} > 0$ for $n = 1, \ldots, k$. Using (1.25) we get

$$0 < \mathcal{L}(\pi_j^2) = \frac{\Delta_j}{\Delta_{j-1}}, \quad \text{for } j = 0, \ldots, k.$$

Recalling that $\Delta_{-1} = 1$ we get $\Delta_j > 0$ for $j = 0, \ldots, k$. □

In addition, in Appendix A we show that $\mathcal{L}$ is positive definite on $\mathcal{P}_k$ if and only if there exists a positive non-decreasing distribution function μ supported on the real axis such that $\mathcal{L}(p) = \int p(x)\mathrm{d}\mu(x)$ for all p from $\mathcal{P}_{2k}$. Hence, the classical theory of orthogonal polynomials concern positive definite linear functionals.

Zeros of the orthogonal polynomials with respect to a positive definite functional have a regular behavior.

Definition 1.13 ([15, Chapter I, Definition 5.1])**.** *Given a subset $E \subset (-\infty, +\infty)$, a linear functional $\mathcal{L}$ is* positive definite on E *if and only if $\mathcal{L}(p) > 0$ for every polynomial p non-negative and not identically zero on E. We say that E is a* supporting set *for $\mathcal{L}$.*

Theorem 1.14 ([15, Chapter I, Theorem 5.1])**.** *If $\mathcal{L}$ is positive definite on a infinite subset $E \subset (-\infty, +\infty)$, then it is positive definite on every set containing E and on every dense subset of E.*

Proof. Let p be a polynomial nonnegative and not identically zero on S. Since E is an infinite set p cannot be identically zero on E. If S is a subset of E, then trivially p is nonnegative, and not identically zero, on E. Hence, $\mathcal{L}(p) > 0$. Conversely, let S be a dense subset of E. By continuity $p(x) \geq 0$ for every $x \in E$ and it is not identically zero. □

Thus, in general, the "smallest" infinite supporting set does not exist for a positive definite linear functional.

The following theorem will be fundamental for the definition of Gauss quadrature and, moreover, to understand the main problem behind the extension of a Gauss-like quadrature for the case of general linear functionals. We refer in particular to [87, Theorem 3.3.1] and [15, Chapter I, Theorem 5.2].

Theorem 1.15. *Let I be an interval and a supporting set for a positive definite functional $\mathcal{L}$ on $\mathcal{P}_k$. Let $p_0, \ldots, p_k$ be orthogonal polynomials with respect to $\mathcal{L}$. The zeros of p_n are all real, simple and located in the interior of I, for $n = 0, \ldots, k$.*

Proof. At least one of the roots of p_n must lay in I. Indeed, $\mathcal{L}(p_n) = 0$ implies that p_n cannot be positive (or negative) on the interior of the supporting set I. Let $\lambda_1, \ldots, \lambda_\ell$ be the distinct zeros of p_n with odd multiplicity that lay in the interior of I. Multiplying $r(x) = (x - \lambda_1) \cdots (x - \lambda_\ell)$ by p_n gives $r(x)p_n(x)$, a polynomial without zeros of odd multiplicity in the interior of I. Thus, since $r(x)p_n(x) \geq 0$ for $x \in I$ we get $\mathcal{L}(rp_n) > 0$. Since p_n is an orthogonal polynomial this is a contradiction unless $\ell \geq n$, hence unless $\ell = n$. Therefore, p_n has n distinct zeros in the interior of I, for $n = 0, \ldots, k$. □

We present the *interlacing property* for the zeros of a sequence of orthogonal polynomial. In this section we prove it using orthogonal polynomials properties. In the literature we can find several different formulations of this result [84, Chapter 1, Section 3], [87, Theorem 3.3.2 and 3.3.3], [15, Chapter I, Theorem 5.3 and Chapter II, Section 4] and [63, Theorem 3.3.1]. Since it is possible to give this proof through many approaches the property has been rediscovered many times, see [53, Chapter 4, Theorem 4, p. 168], [95, Chapter 2, Section 41 and 47], [38, Theorem 2, p. 121] and [89, Theorem 6.1, pp. 663–664].

Theorem 1.16 (Strict Interlacing Property)**.** *Let $\mathcal{L}$ be a linear functional positive definite on $\mathcal{P}_k$ and let $\pi_0, \ldots, \pi_k$ its related monic orthogonal polynomials. If $\lambda_0^{(n)}, \ldots, \lambda_k^{(n)}$ are the zeros of π_n for $n = 0, \ldots, k$, then*

$$\lambda_i^{(n+1)} < \lambda_i^{(n)} < \lambda_{i+1}^{(n+1)}, \quad \text{for } n = 0, \ldots, k-1, \quad i = 1, \ldots, n.$$

Proof. By Theorem 1.15 π_n has real and distinct zeros

$$\lambda_1^{(n)} < \lambda_2^{(n)} < \cdots < \lambda_n^{(n)}, \quad \text{for } n = 0, \ldots, k.$$

Thus the zeros of π_n' are all distinct from the zeros of π_n. Moreover, by Rolle's Theorem, π_n' has at least one zero, and hence exactly one, on each interval $(\lambda_i^{(n)}, \lambda_{i+1}^{(n)})$, for $i = 1, \ldots, n-1$. Again, the zeros of π_n' are real and distinct, hence $\pi_n'(\lambda_i^{(n)})$ alternates its sign as $i = 1, \ldots, n$. The leading coefficient of π_n' is positive thus

$$\pi_n'(x) > 0 \quad \text{for } x \geq \lambda_n^{(n)}, \quad \operatorname{sgn}(\pi_n'(x)) = (-1)^n \quad \text{for } x \leq \lambda_1^{(n)},$$

with sgn the sign function. Therefore

$$\operatorname{sgn}(\pi_n'(\lambda_i^{(n)})) = (-1)^{n-i}, \quad \text{for } i = 1, \ldots, n. \tag{1.26}$$

Since $\mathcal{L}$ is positive definite on $\mathcal{P}_k$ (1.24) gives

$$\pi_{n+1}'(x)\pi_n(x) - \pi_n'(x)\pi_{n+1}(x) > 0$$

for $n = 0, \ldots, k-1$. Therefore

$$\pi'_{n+1}(\lambda_i^{(n+1)})\pi_n(\lambda_i^{(n+1)}) > 0, \quad \text{for } n = 0, \ldots, k-1, \quad i = 1, \ldots, n+1.$$

Using (1.26) we get $\operatorname{sgn}(\pi_n(\lambda_i^{(n+1)})) = (-1)^{n+1-i}$, for $n = 0, \ldots, k-1$, $i = 1, \ldots, n+1$, that finishes the proof. □

Finally, if $\mathcal{L}$ is positive definite on $\mathcal{P}$, we have the following immediate consequence of the interlacing property. For each fixed i the sequence $\lambda_i^{(1)}, \lambda_i^{(2)}, \ldots$ is a decreasing sequence, and the sequence $\lambda_1^{(i)}, \lambda_2^{(i+1)}, \ldots$ is an increasing sequence. Hence, both $\lim_{n\to\infty} \lambda_i^{(n)}$ and $\lim_{n\to\infty} \lambda_n^{(n-i+1)}$ exist in the extended real number line.

CHAPTER 2

Matrix Functions

2.1 Definition and Properties

Since the power of a matrix is a basic concept it is natural to define the *matrix polynomial* as the function $p : \mathbb{C}^{k\times k} \to \mathbb{C}^{k\times k}$ given by

$$p(A) = a_n A^n + a_{n-1}A^{n-1} + \cdots + a_0 I, \quad \text{for } A \in \mathbb{C}^{k\times k}, \tag{2.1}$$

with I the identity matrix. In the same spirit, the idea is to define a function of matrices $f : \mathbb{C}^{k\times k} \to \mathbb{C}^{k\times k}$ not elementwise, but substituting a matrix to the variable of a scalar function. Moreover, if the series $f(x) = \sum_{n=0}^{\infty} a_n x^n$ converges for $x \in C \subset \mathbb{C}$, we would like to define the *matrix function* as

$$f(A) = \sum_{n=0}^{\infty} a_n A^n,$$

with A in a subset of $\mathbb{C}^{k\times k}$ for which the series is convergent. However, this is only one of the equivalent approaches by whom we can define a *matrix function*. Indeed, as noticed by Rinehart in [76], eight equivalent definitions have been given since 1880 by Weyr [93], Sylvester and Buchheim [86, 12], Giorgi [37], Cartan, Fantappiè [28], Cipolla [19], Schwerdtfeger [79] and Richter [75].

To define matrix functions and give their properties we will start from one definition, useful for our case, and then we will show the equivalence of this definition with some other possible definitions. We refer to [51, in particular Chapter 1, Section 2] for a deeper discussion. First, we recall the definition of *Jordan normal form* of a matrix.

Definition 2.1 (Jordan normal form)**.** *Each $k \times k$ matrix A with values in $\mathbb{C}$ is similar to a matrix in the Jordan normal form, i.e., there exist an invertible matrix W such that $W^{-1}AW = \mathrm{diag}(\Lambda_1, \ldots, \Lambda_\nu) = \Lambda$ a block diagonal matrix, with*

$$\Lambda_i = \Lambda_{s_i}(\lambda_i) = \begin{bmatrix} \lambda_i & 1 & 0 & \ldots & 0 \\ 0 & \lambda_i & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & \lambda_i \end{bmatrix} \in \mathbb{C}^{s_i \times s_i}.$$

These matrices are called Jordan blocks*, $s_1 + \cdots + s_\nu = k$ and $\lambda_1, \ldots, \lambda_\nu$ are the eigenvalues of A (not necessarily distinct). Moreover, we call $s(\lambda)$,* index *of the eigenvalue λ, the size of the largest Jordan block associated with λ.*

We recall that the Jordan normal form of a matrix is not unique. However, it is unique up to the order of the Jordan blocks. Naming the columns of W

$$\mathbf{w}_{1,1}, \ldots, \mathbf{w}_{1,s_1}, \mathbf{w}_{2,1}, \ldots, \mathbf{w}_{2,s_2}, \ldots, \mathbf{w}_{\nu,1}, \ldots, \mathbf{w}_{\nu,s_\nu},$$

from $AW = W\Lambda$, we get the relation

$$(A - \lambda_i)\mathbf{w}_{i,j} = \mathbf{w}_{i,j-1}, \quad \mathbf{w}_{i,0} = 0, \quad \text{for } i = 1, \ldots, \nu, \quad j = 1, \ldots, s_i.$$

Clearly, $\mathbf{w}_{i,1}$ is an eigenvector of A associated with the eigenvalue λ_i. Conversely, if W is an invertible matrix whose columns $\mathbf{w}_{i,2}, \ldots, \mathbf{w}_{i,s_i}$ satisfy the previous relation, then $W^{-1}AW = \Lambda$, with Λ a Jordan normal form of A. We recall that vectors $\mathbf{w}_{i,2}, \ldots, \mathbf{w}_{i,s_i}$ satisfying the previous relation are known as *generalized eigenvectors* of A (or *Jordan canonical vectors* of A or *principal vectors*) associated with the eigenvalue λ_i. It will be useful to remember that the number of Jordan blocks associated with the same eigenvalue λ is equal to the dimension of the eigenspace of λ, i.e., the *geometric multiplicity* of λ.

We say that a function f is *defined on the spectrum of the given matrix* J if for every eigenvalue λ_i of J there exists $f^{(j)}(\lambda_i)$ for $j = 0, 1, \ldots, s(\lambda_i)$, (where $s(\lambda_i)$ is the index of λ_i); see, e.g., [51, p. 3, Definition 1.1] We can give the following definition of matrix function, first given in [37], see [51, p. 3, Definition 1.2].

Definition 2.2 (Matrix function)**.** *Let f be a function defined on the spectrum of a given matrix A and $W^{-1}AW = diag(\Lambda_1, \ldots, \Lambda_\nu)$ be the Jordan normal form of A. The matrix function $f(A)$ is then defined as*

$$f(A) = W\, diag(f(\Lambda_1), \ldots, f(\Lambda_\nu))W^{-1},$$

with

$$f(\Lambda_i) = \begin{bmatrix} f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \frac{f^{(2)}(\lambda_i)}{2!} & \dots & \frac{f^{(s_i-1)}(\lambda_i)}{(s_i-1)!} \\ 0 & f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \dots & \frac{f^{(s_i-2)}(\lambda_i)}{(s_i-2)!} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{f'(\lambda_i)}{1!} \\ 0 & \dots & \dots & 0 & f(\lambda_i) \end{bmatrix},$$

for $i = 1, \dots, \nu$.

Naturally, when A is a $k \times k$ diagonalizable matrix its Jordan normal form is a diagonal matrix $W^{-1}AW = \mathrm{diag}(\lambda_1, \dots, \lambda_k)$, with $\lambda_1, \dots, \lambda_n$ the eigenvalues of A. Hence, if f is a function defined on the spectrum of A, then $f(A) = W\mathrm{diag}(f(\lambda_1), \dots, f(\lambda_k))W^{-1}$.

Definition 2.2 seems to depend on the Jordan normal form of the matrix. However, in the following we will show that the definition is independent from the chosen Jordan normal form. For the moment let us assume a choice of a Jordan normal form for the given matrix.

By direct computation we have the following properties.

Lemma 2.3. *Let A be a complex $k \times k$ matrix and f, g be functions defined on the spectrum of A, then*

1. $(f + g)(A) = f(A) + g(A)$;
2. $(f \cdot g)(A) = f(A)g(A)$;
3. *if $f(x) = \alpha x$, with $\alpha \in \mathbb{C}$, then $f(A) = \alpha A$;*
4. *if $f(x) = \alpha \in \mathbb{C}$, then $f(A) = \alpha I_k$;*
5. *if $f(x) = 1/(\alpha - x)$ (α cannot be an eigenvalue of A), then $f(A) = (\alpha I_k - A)^{-1}$, i.e., it is the resolvent of A at α.*

We remark that, given an invertible matrix A, by Property 5 the inverse of the matrix A^{-1} is the matrix function $f(A)$ with $f(x) = 1/x$. Moreover, let $f_j(x) = x^j$ for $j \in \mathbb{Z}$, then properties 2, 4 and 5 of Lemma 2.3 give

$$f_j(A) = A^j, \quad \text{for } j \in \mathbb{Z}$$

for every complex matrix A. Hence, from now on A^j will represent both the (typical) j-th power of a matrix and the matrix function f_j apply to A. Thus, by Property 1 and 3 of Lemma 2.3 we get that given a polynomial $p(x) = a_n x^n + \dots + a_0$ the matrix function $p(A)$ is equal to $a_n A^n + \dots + a_0 I_k$

for every $k \times k$ complex matrix A. This shows that a polynomial matrix function is equivalent to the matrix polynomial (2.1).

Another equivalent way to define matrix functions $f(A)$ is through generalized Hermite interpolation (see, e.g., [51, p. 5, Theorem 1.3]). We say that a polynomial p interpolate f on the spectrum of A in the Hermite sense if

$$p(\lambda_i)^{(j)} = f(\lambda_i)^{(j)}, \quad \text{for } i = 1, \dots, \ell, \quad j = 0, \dots, s(\lambda_i) - 1,$$

with λ_i the eigenvalue of A of index $s(\lambda_i)$.

Corollary 2.4. *A polynomial p interpolates a function f on the spectrum of a matrix A in the Hermite sense if and only if $f(A) = p(A)$. Moreover, there exists a unique polynomial p such that $f(A) = p(A)$ with degree lower than or equal to the degree of the minimal polynomial of A.*

Proof. If p interpolates f on the spectrum of a matrix A in the Hermite sense, then by Definition 2.2 $p(A) = f(A)$. Vice versa, if $f(A) = p(A)$, then

$$f(A) = W f(\Lambda) W^{-1} = W p(\Lambda) W^{-1} = p(A),$$

with $\Lambda = W^{-1}AW$ the Jordan normal form of A. Since $f(\Lambda) = p(\Lambda)$ p interpolates f on the spectrum of a matrix A in the Hermite sense.

Finally, there exists a polynomial p satisfying the conditions

$$p(\lambda_i)^{(j)} = f(\lambda_i)^{(j)}, \quad \text{for } i = 1, \dots, \ell, \quad j = 0, \dots, s(\lambda_i) - 1,$$

with λ_i the eigenvalue of A of index $s(\lambda_i)$. The minimal degree of p is lower than or equal to $s(\lambda_1) + \cdots + s(\lambda_\ell)$, that is, the degree of the minimal polynomial of A. Moreover, p is unique. □

Corollary 2.4 then shows that the definition of $f(A)$ is independent from the chosen Jordan normal form in Definition 2.2.

It is important to remark that the polynomial p of Corollary 2.4 depends on the matrix A. However, given two matrices A, B it is always possible to define a polynomial p such that $f(A) = p(A)$ and $f(B) = p(B)$. Indeed, it is enough to choose a polynomial that interpolates f on the spectrum of A and B in the Hermite sense. Naturally, the degree of p will be lower than or equal to the sum of the degrees of the minimal polynomials respectively of A and B.

The equivalence given by Corollary 2.4 allows us to prove the following properties for matrix functions. First, we remark that if p is a polynomial, then it is defined on the spectrum of every complex matrix A. Moreover,

$$p(A^*) = (\bar{p}(A))^*, \tag{2.2}$$

with $\bar{p}$ the polynomial whose coefficients are the conjugate coefficients of p. Indeed,

$$\begin{aligned} p(A^*) &= \alpha_n (A^*)^n + \alpha_{n-1}(A^*)^{n-1} \cdots + \alpha_0 I \\ &= \alpha_n (A^n)^* + \alpha_{n-1}(A^{n-1})^* \cdots + \alpha_0 I = (\bar{p}(A))^*. \end{aligned}$$

Proposition 2.5 (see [51, p. 13, Theorem 1.18])**.** *If the function f is defined on the spectrum of the matrix A, then $f(A^*) = (f(A))^*$ if and only if*

$$f^{(j)}(\bar{\lambda}_i) = \overline{f^{(j)}(\lambda_i)}, \quad \text{for } i = 1, \ldots, \ell, \quad j = 0, \ldots, s(\lambda_i), \tag{2.3}$$

with λ_i eigenvalues with index $s(\lambda_i)$.

Proof. By Corollary 2.4 there exist a polynomial p such that $f(A) = p(A)$ and $f(A^*) = p(A^*)$. Equation (2.2) then gives

$$f(A^*) = p(A^*) = (\bar{p}(A))^*.$$

By Definition 2.2 $\bar{p}(A) = f(A) = p(A)$ if and only if (2.3) is satisfied. □

Proposition 2.6. *Let A be a matrix and X be an invertible matrix. Moreover, let the function f be defined on the spectrum of A and XAX^{-1}. Then*

$$f(XAX^{-1}) = Xf(A)X^{-1}.$$

Proof. By Corollary 2.4 there exists a polynomial $p(x) = \alpha_n x^n + \ldots \alpha_0$ for which $f(A) = p(A)$ and $f(XAX^{-1}) = p(XAX^{-1})$. Thus

$$f(XAX^{-1}) = \alpha_n XA^nX^{-1} + \ldots \alpha_0 I = Xp(A)X^{-1} = Xf(A)X^{-1}.$$

□

With a similar proof we can show that if X permutes with A, then it also permutes with $f(A)$.

In 1928 E. Cartan proposed in a letter to G. Giorgi to define matrix functions using the Cauchy integral formula, see [76, Section 2.3]. The following proposition gives the equivalence of this definition with Definition 2.2, when f is an analytic function.

Proposition 2.7. *Let f be an analytic function on some open $\Omega \in \mathbb{C}$, and $\Gamma \in \Omega$ be a system of Jordan curves encircling each eigenvalue of A exactly one time, with mathematical positive orientation, then*

$$f(A) = \int_\Gamma f(x) \left(xI - A\right)^{-1} dx.$$

Proof. The equivalence was first given in [76]. By Property 5 of Lemma 2.3 $(xI-A)^{-1}$ is a matrix function. Hence, it is enough to show that this formula stands for a Jordan block Λ_i, i.e.,

$$(xI-\Lambda_i)^{-1} = \begin{bmatrix} \frac{1}{(x-\lambda_i)} & \frac{1}{(x-\lambda_i)^2} & \cdots & \frac{1}{(x-\lambda_i)^{s_i}} \\ 0 & \frac{1}{(x-\lambda_i)} & \cdots & \vdots \\ \vdots & \ddots & \ddots & \frac{1}{(x-\lambda_i)^2} \\ 0 & \cdots & \cdots & \frac{1}{(x-\lambda_i)} \end{bmatrix}.$$

Using Cauchy integral formula for f and its derivative elementwise finishes the proof. □

Consider now a function f analytic at $x_0 \in \mathbb{C}$, then we have

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \cdots + \frac{f^n(x_0)}{n!}(x-x_0)^n + \ldots. \tag{2.4}$$

We say that a sequence of matrices $A_0, A_1, A_2, \ldots$ converges if it converges elementwise, i.e., each sequence given by the corresponding elements is convergent. Moreover, we say that an infinite series of matrix $\sum_{i=0}^{\infty} A_i$ is convergent if the sequence of the partial sums converges. We want to show that the matrix obtained by a convergent series

$$f(x_0)I + f'(x_0)(A-x_0I) + \cdots + \frac{f^n(x_0)}{n!}(A-x_0I)^n + \ldots \tag{2.5}$$

is equal to the matrix function $f(A)$. We remark that this is the way in which Weyr defined a matrix function in [93]. The following theorem for the convergence of series (2.5) was first proved by Hensel in [47] for Maclaurin series (see also [76]).

Theorem 2.8. *The power series (2.5) converges if and only if every eigenvalue of A lies within or on the circle of convergence of the series $f(z)$ (2.4).*

We will show this in the proof of the following proposition.

Proposition 2.9 (see [76, pp. 12–13])**.** *Let the series (2.5) be convergent, then it is equal to the matrix function $f(A)$, with f given by the corresponding scalar series (2.4).*

Proof. Let us consider the partial sum

$$S_n(A) = f(x_0)I + f'(x_0)(A-x_0I) + \cdots + \frac{f^n(x_0)}{n!}(A-x_0I)^n.$$

$S_n(A)$ is a matrix polynomial for every $n = 0, 1, \ldots$, hence, it can be rewritten using Definition 2.2 as $S_n(A) = WS_n(\Lambda)W^{-1}$, with $\Lambda = W^{-1}AW$ the Jordan normal form of A. Therefore, it is enough to show that $S_n(\Lambda_i)$ converges to $f(\Lambda_i)$, for every Λ_i Jordan block of Λ of dimension s_i associated with the eigenvalue λ_i. Indeed,

$$S_n(\Lambda_i) = \begin{bmatrix} S_n(\lambda_i) & S_n'(\lambda_i) & \ldots & \frac{S_n^{(s_i-1)}(\lambda_i)}{(s_i-1)!} \\ 0 & S_n(\lambda_i) & \ldots & \vdots \\ \vdots & \ddots & \ddots & S_n'(\lambda_i) \\ 0 & \ldots & \ldots & S_n(\lambda_i) \end{bmatrix}.$$

Then, $S_n(A)$ converges if and only if the sequence $S_n(\lambda_i)$ and its derivatives sequences $\frac{S_n^{(j)}(\lambda_i)}{(s_i-1)!}$ converge, for $j = 1, \ldots, s_i - 1$. Since the derivative of a power series has the same radius of convergence of the series itself, $S_n(A)$ converges if and only if every eigenvalue of A lies within or on the circle of convergence of the series $f(z)$ (2.4), which proves Theorem 2.8. Moreover, if (2.5) converges, then it converges to $f(A)$. □

Matrix exponential.
We conclude this chapter introducing one of the most well-known matrix function: the matrix exponential; we refer to [51, Chapter 10]. Many mathematical models for physical, biological, and economic problems, in particular the solution of parabolic PDE equations (such as the heat equation), involve the solution of ordinary differential equations

$$\dot{z}(t) = Az(t), \quad t \in \mathbb{R}$$

with A a square complex matrix and $z(0) = z_0$ the initial condition. The solution is given by

$$z(t) = e^{At}z_0 = \exp(At)z_0,$$

with $\exp(At)$ the matrix exponential obtained by Definition 2.2. Moreover, in Chapter 6 we will see an application of the matrix function $\exp(A)$ in the *complex networks theory*.

The matrix function $\exp(A)$ is defined for every complex matrix A, indeed, the exponential is defined on the spectrum of every complex matrix A, or equivalently, the Maclaurin series for the exponential

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$$

is convergent for every $x \in \mathbb{C}$.

We recall here some well-known properties of the matrix exponential.

Proposition 2.10. *Let A, B be complex matrices of the same dimension, then we have the following properties for the matrix exponential:*

1. $\frac{d}{dx}e^{Ax} = Ae^{Ax}$, *for* $x \in \mathbb{C}$;
2. $e^{\mathbf{0}} = I$, *for every null matrix* $\mathbf{0}$;
3. $e^{At}e^{As} = e^{A(t+s)}$, *for* $s, t \in \mathbb{R}$;
4. *If* $AB = BA$, *then* $e^{At}e^{Bt} = e^{(A+B)t}$, *for* $t \in \mathbb{R}$;
5. $(e^A)^{-1} = e^{-A}$

Proof. 1. Using the series form of the derivative of $\exp(Ax)$ gives

$$\frac{\mathrm{d}}{\mathrm{d}x}e^{Ax} = \frac{\mathrm{d}}{\mathrm{d}x}\sum_{i=0}^{\infty}\frac{A^i x^i}{i!}.$$

Considering the partial sum up to n we get

$$\frac{\mathrm{d}}{\mathrm{d}x}\sum_{i=0}^{n}\frac{A^i x^i}{i!} = \frac{\mathrm{d}}{\mathrm{d}x}I + \sum_{i=1}^{n}\frac{\mathrm{d}}{\mathrm{d}x}\frac{A^i x^i}{i!} = \sum_{i=1}^{n}\frac{A^i x^{i-1}}{(i-1)!} = A\sum_{j=0}^{n-1}\frac{A^j x^j}{j!}.$$

Letting $n \to \infty$ finishes the proof.

2. It directly comes from Definition 2.2.

3. Let $\mathbf{y}(t) = \left(e^{At}e^{As} - e^{A(t+s)}\right)\mathbf{y}_0$, then by Property 1

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}(t) = (Ae^{At}e^{As} - Ae^{A(t+s)})\mathbf{y}_0 = A\mathbf{y}(t),$$

for every $\mathbf{y}_0 \in \mathbf{C}$. Hence, Solving the previous differential equation we get $\mathbf{y}(t) = e^{At}\mathbf{y}(0)$. Since $\mathbf{y}(0) = 0$ we get $\mathbf{y}(t) \equiv 0$, which concludes the proof.

4. By Corollary 2.4 we can write e^{At} as a polynomial. Hence, since A and B commute

$$e^{At}B = \left(\sum_{i=0}^{\ell}\alpha_i A^i t^i\right)B = Be^{At}.$$

Now, let us define $\mathbf{y}(t) = \left(e^{At}e^{Bt} - e^{(A+B)t}\right)\mathbf{y}_0$, then

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathbf{y}(t) = (Ae^{At}e^{Bt} + e^{At}Be^{Bt} - (A+B)e^{(A+B)t})\mathbf{y}_0 = (A+B)\mathbf{y}(t).$$

We finish using the same arguments as in the previous proof.

5. From properties 2 and 3 we obtained $e^A e^{-A} = e^{\mathbf{0}} = I$. □

We remark that by Property 3 of Proposition 2.10

$$\left(e^{(A/m)}\right)^m = e^A,$$

for every complex A and $m = 1, 2, \ldots$.

However, not all the properties we would like to have are satisfied by the exponential. For example, let

$$A_1 = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Since $A_1A_2 \neq A_2A_1$, Property 4 of Proposition 2.10 may not hold. Indeed, we get

$$e^{A_1} = \begin{bmatrix} e^2 & 0 \\ 0 & e \end{bmatrix}, \quad e^{A_2} = \begin{bmatrix} e & e \\ 0 & e \end{bmatrix}$$

and so

$$e^{A_1}e^{A_2} = \begin{bmatrix} e^3 & e^3 \\ 0 & e^2 \end{bmatrix} \neq e^{A_1+A_2} = \begin{bmatrix} e^3 & e \\ 0 & e^2 \end{bmatrix}.$$

In this chapter we showed that matrix functions from Definition 2.2 satisfy many properties that, intuitively, we would like a matrix function to have. Indeed, the definition comprehends polynomials of matrices (2.1), satisfies the Cauchy integral formula from Proposition 2.7 for analytic functions, and, when we have convergence, it is equivalent to use a matrix instead of the scalar variable in the Taylor series of a function. Moreover, it has the basic properties of Lemma 2.3. However, we must be careful, since important properties of specific scalar functions are not true for the corresponding matrix functions, as we have seen in the previous examples.

CHAPTER 3

Jacobi Matrices

3.1 Definition

In Chapter 1 we saw that any sequence of orthonormal polynomials $p_0, \ldots, p_{n-1}$ is associated with a tridiagonal symmetric matrix (1.10) with nonzero elements on its sub- and super-diagonal. Let $\mathcal{L}$ be a positive definite linear functional (see Definition 1.10), then Definition 1.10, (1.12) and (1.13) show that the polynomials $p_0, \ldots, p_{n-1}$ orthonormal with respect to $\mathcal{L}$ are real and, in addition, the tridiagonal symmetric matrix (1.10) is a real matrix. Usually a real tridiagonal symmetric matrix with nonzero elements on its sub- and super-diagonal is called a *Jacobi matrix*. However, we are dealing with quasi-definite linear functionals (Definition 1.2). Hence, the matrix (1.10) may be complex.

There are many different definitions of Jacobi matrices in the literature. In the most frequent one, a Jacobi matrix is defined as a real, symmetric, tridiagonal matrix with positive elements on the super-diagonal ([1, p. 2], [18, p. 72], [42, p. 13], [63, p. 30]). Jacobi matrices are important in matrix computations (approximating eigenvalues and eigenvectors or solving linear algebraic systems) and in approximation theory (approximating functions and integrals). They were named after Carl Gustav Jacob Jacobi (1804-1851), one of the most prolific mathematician of the 19th century. He proved that using a linear transformation with determinant equal to ± 1 it is possible to reduce any quadratic form with n variables into a particular quadratic form defined by $2n - 1$ coefficients (see [55]). Nowadays, this last quadratic form

can be expressed in terms of the $n \times n$ Jacobi matrix. To our knowledge, the first Jacobi matrix appeared on [49, p. 202]. A paper in which Toeplitz and Hellinger discussed the relationship between quadratic forms with infinitely many unknowns and the analytic theory of continued fractions by Stieltjes [84]. For a detailed history of Jacobi matrices we refer to [63, Section 3.4.3].

Other definitions of Jacobi matrices can be found in [32, Vol. 2, p. 99] (a real tridiagonal matrix), [52, p. 86] (a tridiagonal matrix with a real diagonal and such that the product of the corresponding elements of the sub- and super-diagonal is non-negative), [50] (a tridiagonal symmetric matrix with a complex diagonal and with nonzero real elements on the sub- and super-diagonal). In this paper we use the definition by Beckermann from the paper about spectral properties of *complex Jacobi matrices* [2].

Definition 3.1 (Jacobi matrix)**.** *A square complex matrix is called a* Jacobi matrix *if it is tridiagonal, symmetric and has no zero elements on its sub- and super-diagonal.*

Probably the first study of this class of matrices appeared in [92, p. 226], where Wall investigated the convergence of complex Jacobi continued fractions (J-fractions). We remark that a (complex) Jacobi matrix is Hermitian if and only if it is real.

As described in Section 1.1, k orthonormal polynomials $p_0, p_1, \ldots, p_{k-1}$, determine a Jacobi matrix J_k. Conversely, by Favard Theorem 1.5 every Jacobi matrix defines a sequence of polynomials orthogonal with respect to a certain linear functional. Moreover, if there exist $n-1$ orthogonal polynomials, then we have n Jacobi matrices $J_1, \ldots, J_n$. Since adding a polynomial to a set of orthogonal polynomials does not change the three-term relationship among the original set, then J_{k-1} is the $k-1 \times k-1$ leading principal submatrix of J_k for $k = 2, \ldots, n$. Let us show it with an example.

Example 3.1 Let $\mathcal{L}$ be a linear functional defined by a sequence of moments with the first seven terms given by

$$1,\ 3,\ 8,\ 20,\ 52,\ 156,\ i.$$

Please, notice that here $i = \sqrt{-1}$ is the imaginary unit. Then $\mathcal{L}$ is quasi-definite on $\mathcal{P}_3$, since

$$\Delta_0 = 1, \quad \Delta_1 = -1, \quad \Delta_2 = -4, \quad \Delta_3 = 2128 - 4i.$$

The associated monic orthogonal polynomials are

$$\pi_0 = 1, \quad \pi_1(x) = x - 3, \quad \pi_2(x) = x^2 - 4x + 4, \quad \pi_3(x) = x^3 - 7x^2 + 20x - 24,$$

and a sequence of orthonormal polynomials is

$$p_0 = 1, p_1(x) = \frac{x-3}{i}, p_2(x) = -\frac{x^2-4x+4}{2}, p_3(x) = \frac{2x^3-14x^2+40x-48}{\sqrt{i-532}}.$$

Then

$$J_1 = [\,3\,], \quad J_2 = \begin{bmatrix} 3 & i \\ i & 1 \end{bmatrix}, \quad J_3 = \begin{bmatrix} 3 & i & 0 \\ i & 1 & 2i \\ 0 & 2i & 3 \end{bmatrix}$$

are the corresponding first 3 Jacobi matrices.

In this Chapter we will show some spectral properties of Jacobi matrices. We will first see some theorems about complex tridiagonal matrices (Section 3.2), then we will investigate complex symmetric matrices (Section 3.3). In Section 3.4 we will prove the moment matching property for quasi-definite linear functionals and complex Jacobi matrices. Finally, in Section 3.5 we will recall some additional properties of real Jacobi matrices.

3.2 Complex Tridiagonal Matrices

Some spectral properties of complex tridiagonal matrices (with nonzero elements on the sub- and super-diagonal) will be important for the following chapters. We first recall the main results.

Theorem 3.2. *Every tridiagonal matrix $T \in \mathbb{C}^{n\times n}$ with nonzero elements on its super-diagonal (or sub-diagonal) is non-derogatory, i.e., its eigenvalues have geometric multiplicity* 1*.*

Proof. Let λ be an eigenvalue of a tridiagonal matrix T with nonzero elements on the super-diagonal (the other case is analogous). Deleting the first column and the last row of $T-\lambda I$ gives a lower triangular non-singular matrix. Thus, the null space of $T-\lambda I$ has dimension 1 because its rank is not smaller than $n-1$. □

Corollary 3.3. *Every tridiagonal matrix $T \in \mathbb{C}^{n\times n}$ with nonzero elements on its super-diagonal (or sub-diagonal) is diagonalizable if and only if it has distinct eigenvalues.*

It is known that we can use the adjoint matrix (sometimes the term adjugate is used to avoid confusion with the Hermitian adjoint) in order to give an explicit formulation of the eigenvectors corresponding to the eigenvalues of geometric multiplicity one. Indeed, if λ is an eigenvalue with geometric multiplicity one, then $\operatorname{rank}(A-\lambda I) = n-1$. Hence, $\operatorname{adj}(A-\lambda I)$ is not

identically zero. Then let $\mathrm{adj}(A-\lambda I)\mathbf{e}_i$ be a nonzero column of $\mathrm{adj}(A-\lambda I)$. For later convenience, let us consider any $\xi \in \mathbb{C}$, we get

$$(A-\xi I)\,\mathrm{adj}(A-\xi I) = \det(A-\xi I)\,I,$$

and thus the i-th column gives

$$(A-\xi I)\mathbf{z}(\xi) = \det(A-\xi I)\mathbf{e}_i,$$

with $\mathbf{z}(\xi) = \mathrm{adj}(A-\xi I)\mathbf{e}_i$. Fixing $\xi = \lambda$ gives $(A-\lambda I)\mathbf{z}(\lambda) = \mathbf{0}$ which shows that $\mathbf{z}(\lambda)$ is an eigenvector of A associated with λ. We remark that the same eigenvector (apart from the normalization) is given by any nonzero column of $\mathrm{adj}(A-\lambda I)$.

Following [21], we differentiate j times $(A-\xi I)\mathbf{z}(\xi) = \det(A-\xi I)\mathbf{e}_i$, which gives

$$(A-\xi I)\mathbf{z}^{(j)}(\xi) = j\mathbf{z}^{(j-1)}(\xi) + \frac{d^j}{d\xi^j}\det(A-\xi I)\mathbf{e}_i.$$

Denoting

$$\mathbf{w}_0(\xi) = \mathbf{0},\ \ \mathbf{w}_1(\xi) = \mathbf{z}(\xi),\ \ \mathbf{w}_{j+1}(\xi) = \frac{1}{j}\mathbf{w}_j'(\xi) = \frac{1}{j!}\mathbf{z}^{(j)}(\xi), \tag{3.1}$$

for $j = 1, 2, \ldots$, we obtain

$$(A-\xi I)\mathbf{w}_{j+1}(\xi) = \mathbf{w}_j(\xi) + \frac{1}{j!}\frac{d^j}{d\xi^j}\det(A-\xi I)\mathbf{e}_i,\ \text{where } j = 0, 1, \ldots.$$

If λ is an eigenvalue with geometric multiplicity 1 and algebraic multiplicity s, then we get

$$(A-\lambda I)\mathbf{w}_{j+1}(\lambda) = \mathbf{w}_j(\lambda) \text{ for } j = 0, \ldots, s-1. \tag{3.2}$$

Therefore $\mathbf{w}_1(\lambda)$ is the eigenvector and $\mathbf{w}_j(\lambda)$ for $j = 2, \ldots, s$ are the *generalized eigenvectors* of A (Jordan canonical vectors of A) corresponding to λ; see in particular Definition 2.1. Moreover, $\mathbf{w}_1(\lambda) \neq \mathbf{0}$ and (3.2) imply that $\mathbf{w}_2(\lambda), \ldots, \mathbf{w}_s(\lambda)$ are also nonzero vectors.

Let T_n be a tridiagonal matrix of dimension $n \times n$, then $\mathbf{z}(\xi) = \mathrm{adj}(T_n - \xi I)\mathbf{e}_n \neq 0$. By direct computation we then get an explicit formulation

$$\mathbf{z}(\xi) = \begin{bmatrix} \beta_1 \cdots \beta_{n-1} \\ -\beta_2 \cdots \beta_{n-1}\phi_1(\xi) \\ \vdots \\ (-1)^{n-2}\beta_{n-1}\phi_{n-2}(\xi) \\ (-1)^{n-1}\phi_{n-1}(\xi) \end{bmatrix}, \tag{3.3}$$

where $\beta_1, \ldots, \beta_{n-1}$ are the elements of the super-diagonal and $\phi_i(\xi) = \det(T_i - \xi I)$, with T_i the i-th leading principal submatrix of T_n. This was proved for Hermitian tridiagonal matrices by Wilkinson in [95, Chapter 5, Section 48]. Providing that T_n has no zeros on its super- and sub-diagonal, $\phi_i(\xi) = (-1)^i \pi_i(\xi)$, $i = 1, \ldots, n$, where $\pi_1, \ldots, \pi_n$ is the sequence of monic orthogonal polynomials corresponding to T_n. The following property was presented in the lecture of Ipsen at the ILAS 2005 conference.

Proposition 3.4. *Let $T_n \in \mathbb{C}^{n\times n}$ be a tridiagonal matrix with nonzero elements on its super-diagonal. Let λ be an eigenvalue of algebraic multiplicity s and $\mathbf{w}_{j+1}(\lambda)$, for $j = 1, \ldots, s-1$, the corresponding generalized eigenvectors satisfying $(T_n - \lambda I)\mathbf{w}_{j+1}(\lambda) = \mathbf{w}_j(\lambda)$ (with $\mathbf{w}_0 = \mathbf{0}$, $\mathbf{w}_1 = \mathbf{z}(\lambda)$ from (3.3)). Then we can give the following explicit formulation*

$$\mathbf{w}_j(\lambda) = \frac{1}{(j-1)!} \begin{bmatrix} \mathbf{0}_{j-1} \\ \beta_j \cdots \beta_{n-1} \\ (-1)^j \beta_{j+1} \cdots \beta_{n-1} \phi_j^{(j-1)}(\lambda) \\ \vdots \\ (-1)^{n-2} \beta_{n-1} \phi_{n-2}^{(j-1)}(\lambda) \\ (-1)^{n-1} \phi_{n-1}^{(j-1)}(\lambda) \end{bmatrix}, \quad j = 2, \ldots, s,$$

where $\mathbf{0}_\ell$ is the zero vector of length ℓ, $\beta_1, \ldots, \beta_{n-1}$ are the elements of the super-diagonal of T_n and $\phi_i(\lambda) = det(T_i - \lambda I)$, with T_i the i-th leading principal submatrix of T_n.

Proof. Since $\beta_1, \beta_2, \ldots, \beta_{n-1} \neq 0$, every eigenvector of T_n corresponding to the eigenvalue λ can be expressed as a nonzero multiple of $\mathbf{z}(\lambda)$ from (3.3). Using (3.1) we obtain the form of $\mathbf{w}_j(\lambda)$ in the statement. □

Moreover, we can give the following result about eigenvectors.

Proposition 3.5. *Let A a tridiagonal matrix with nonzero elements on the super- and sub-diagonal. Then, the first and the last component of every eigenvector of A is nonzero.*

Proof. The formula (3.3) for $\mathbf{z}(\lambda)$ shows that the first component of an eigenvector is nonzero. In order to prove the same for the last element of an eigenvector, we need to prove $\phi_{n-1}(\lambda) \neq 0$, i.e., that the eigenvalues of T_n and T_{n-1} are distinct. Using a standard argument, if λ is a root of both the orthogonal polynomials ϕ_n and ϕ_{n-1}, then by (1.12) it is also a root of ϕ_{n-2}. Hence, by induction, $\phi_0 = 0$, which is a contradiction. □

3.3 Complex Symmetric Matrices

Here we introduce and prove some spectral properties related to complex symmetric matrices. In the end of the section we will use these results to prove sufficient and necessary conditions for the diagonalizability of a complex Jacobi matrix.

It is important to remark that unlike real symmetric matrices, complex symmetric matrices may not be diagonalizable. This fact is related with the existence (in the complex field) of *isotropic* vectors. An *isotropic* vector is a vector $\mathbf{x}$ such that $\mathbf{x}^T\mathbf{x} = 0$ and $\mathbf{x} \neq 0$ (for example $(1, i)^T$). In the following we present some results proved by Craven in [20, Theorem 3].

Lemma 3.6 ([20], Lemma 5)**.** *Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors in $\mathbb{C}^k$, with $n < k$, such that $\mathbf{v}_i^T\mathbf{v}_j = 0$ for $i \neq j$ and $\mathbf{v}_i^T\mathbf{v}_i = 1$. Then, there exists a vector $\mathbf{v}_{n+1}$ for which $\mathbf{v}_{n+1}^T\mathbf{v}_{n+1} = 1$ and $\mathbf{v}_{n+1}^T\mathbf{v}_i = 0$ for $i = 1, \ldots, n$.*

Proof. The vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent. If we consider the canonical basis $\mathbf{e}_1, \ldots, \mathbf{e}_k$, up to a renumbering of the vectors, we can assume that $\mathbf{e}_1, \ldots, \mathbf{e}_\ell$ are linearly dependent on $\mathbf{v}_1, \ldots, \mathbf{v}_n$, with $\ell \leq n$. Now, let $j = \ell + 1, \ldots, k$, then we can define

$$\mathbf{u}_j = \mathbf{e}_j - \sum_{i=1}^{n} \alpha_i \mathbf{v}_i, \quad \text{with} \quad \alpha_i = \mathbf{v}_i^T \mathbf{e}_j.$$

Direct computation shows that $\mathbf{u}_j^T\mathbf{v}_i = 0$ for $i = 0, \ldots, n$ and $j = \ell+1, \ldots, k$. To end the proof we need to show that there exist $\mathbf{u}_j$ such that $\mathbf{u}_j^T\mathbf{u}_j \neq 0$, indeed we can obtain the vector $\mathbf{v}_{n+1}$ rescaling $\mathbf{u}_j$. First consider

$$\begin{aligned}
\mathbf{u}_j^T\mathbf{u}_j &= \left(\mathbf{e}_j - \sum_{i=1}^{n} \alpha_i \mathbf{v}_i\right)^T \left(\mathbf{e}_j - \sum_{i=1}^{n} \alpha_i \mathbf{v}_i\right) \\
&= 1 - 2\sum_{i=1}^{n} \alpha_i \mathbf{e}_j^T \mathbf{v}_i + \sum_{i=1}^{n} \alpha_i^2 \\
&= 1 - \sum_{i=1}^{n} (\mathbf{e}_j^T \mathbf{v}_i)^2.
\end{aligned}$$

Moreover, $\mathbf{e}_1, \ldots, \mathbf{e}_\ell$ are linearly dependent on $\mathbf{v}_1, \ldots, \mathbf{v}_n$, thus we can rewrite $\mathbf{e}_j$ as

$$\mathbf{e}_j = \sum_{i=1}^{n} \beta_{i,j} \mathbf{v}_i, \quad \text{with} \quad \beta_{i,j} = \mathbf{v}_i^T \mathbf{e}_j, \quad \text{for} \quad j = 1, \ldots, \ell.$$

Then,

$$1 = \mathbf{e}_j^T \mathbf{e}_j = \sum_{i=1}^{n} (\beta_{i,j})^2 = \sum_{i=1}^{n} (\mathbf{v}_i^T \mathbf{e}_j)^2.$$

By contradiction, assuming $\mathbf{u}_j^T \mathbf{u}_j = 0$ for $j = \ell + 1, \ldots, n-1$, then

$$\sum_{i=1}^{n} (\mathbf{e}_j^T \mathbf{v}_i)^2 = 1, \quad \text{for} \quad j = \ell + 1, \ldots, k.$$

Hence

$$k = \sum_{j=1}^{k} \left(\sum_{i=1}^{n} (\mathbf{e}_j^T \mathbf{v}_i)^2 \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{k} (\mathbf{e}_j^T \mathbf{v}_i)^2 \right) = \sum_{i=1}^{n} \mathbf{v}_i^T \mathbf{v}_i = n.$$

Since we assumed $n < k$ we have a contradiction. □

By induction on the result of Lemma 3.6 we have the following theorem.

Theorem 3.7 ([20], Theorem 2)**.** *Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be vectors as in Lemma 3.6, then there exist vectors $\mathbf{v}_{n+1}, \ldots, \mathbf{v}_k$ such that*

$$\mathbf{v}_i^T \mathbf{v}_j = 0, \text{ if } i \neq j, \quad \text{and} \quad \mathbf{v}_i^T \mathbf{v}_i = 1,$$

for $i, j = 1, \ldots, k$.

Using the previous statements we can give the following result.

Theorem 3.8 ([20], Theorem 3)**.** *If A is a complex $k \times k$ symmetric matrix, then the following statements are equivalent:*

1. *There exists a (complex) nonsingular matrix V such that $V^{-1} = V^T$ and $V^T A V$ is a diagonal matrix;*

2. *Every eigenspace of A has a basis $\mathbf{v}_1, \ldots, \mathbf{v}_s$ without isotropic vectors and such that $\mathbf{v}_i^T \mathbf{v}_j = 0$ for $i \neq j$.*

Proof. The implication 1. $\Rightarrow$ 2. is trivial. Hence, it remains to show the opposite implication. If the union of the eigenspaces bases of the kind of 2. is a basis for $\mathbb{C}^k$, then A is diagonalizable and we get the first statement. If not, then we have a basis $\mathbf{w}_1, \ldots, \mathbf{w}_n$, with $n < k$ satisfying conditions

$$\mathbf{w}_i^T \mathbf{w}_j = 0, \quad \text{for } i \neq j \quad \text{and } \mathbf{w}_i^T \mathbf{w}_i = 1. \tag{3.4}$$

By Theorem 3.7 then we can complete this basis with $\mathbf{w}_{n+1}, \ldots, \mathbf{w}_k$, vectors satisfying conditions (3.4). Then, defining the matrix $W = [\mathbf{w}_1, \ldots, \mathbf{w}_k]$ we get

$$W^T A W = \begin{bmatrix} B & \mathbf{0} \\ \mathbf{0} & C \end{bmatrix},$$

with B an $n \times n$ diagonal matrix and $C = [\mathbf{w}_{n+1}, \ldots, \mathbf{w}_k]^T A [\mathbf{w}_{n+1}, \ldots, \mathbf{w}_k]$ a symmetric complex matrix. Naturally, the spectrum of A is given by the union of the spectrum of B and the spectrum of C. Now, take an eigenvalue λ of C and one related eigenvector $\mathbf{u}$. Then A has the same eigenvalue and $V\,\hat{\mathbf{u}}$ is a related eigenvector, with $\hat{\mathbf{u}}^T = [\,\mathbf{0}_n, \mathbf{u}^T\,]$. The eigenvector $V\,\hat{\mathbf{u}}$ is linearly independent of $\mathbf{w}_1, \ldots, \mathbf{w}_n$, indeed, if

$$\sum_{i=1}^{n} \alpha_i \mathbf{w}_i + \beta V\,\hat{\mathbf{u}} = 0,$$

then for $j = 1, \ldots, n$

$$\sum_{i=1}^{n} \alpha_i \mathbf{w}_j^T \mathbf{w}_i + \beta\,(\mathbf{w}_j^T V)\,\hat{\mathbf{u}} = 0.$$

Since $\mathbf{w}_j^T V = 0$, $\alpha_i = 0$ for $i = 1, \ldots, n$ and $\beta = 0$. Therefore, the eigenspace of λ is not span by any subset of $\mathbf{w}_1, \ldots, \mathbf{w}_n$. This contradicts the fact that the union of the bases of the eigenspaces in statement 2. is not a basis for $\mathbb{C}^k$. □

Now, we consider the propositions presented in [80].

Theorem 3.9 ([80], Theorem 1)**.** *If the null space of a complex symmetric matrix A contains a nonzero isotropic vector, then the trace of the adjoint (or adjugate) of the matrix vanishes, i.e., $tr(A) = 0$.*

Proof. Since the null space of A contains a nonzero vector $\det(A) = 0$. Then the dimension of the null space, $\nu(A)$, is greater than zero. If $\nu(A) > 1$ then $\mathrm{adj}(A)$, the adjoint of A, vanishes.

Hence, it remains to prove the theorem for $\nu(A) = 1$. Let $\mathbf{v}$ be the isotropic vector generating the null space. We recall the well-known property

$$A\,\mathrm{adj}(A) = \det(A) I,$$

that gives

$$A\,\mathrm{adj}(A) = \mathbf{0}.$$

This means that each column of $\mathrm{adj}(A)$ is a vector lying on the null space of A, i.e., each column is a scalar multiple of $\mathbf{v}$. Hence, there exists a vector $\mathbf{u}$, for which

$$\mathrm{adj}(A) = \mathbf{v} \cdot \mathbf{u}^T. \tag{3.5}$$

Indeed, it is enough to define $\mathbf{u}$ so that the i-th column of $\mathrm{adj}(A)$ is equal to $\mathbf{u}_i\mathbf{v}$, with $\mathbf{u}_i$ the i-th element of $\mathbf{u}$. By the symmetry of A we have

$$\mathbf{v} \cdot \mathbf{u}^T = \mathbf{u} \cdot \mathbf{v}^T.$$

Multiplying the previous equation by $\mathbf{v}$ gives

$$\mathbf{v} \cdot \mathbf{u}^T\mathbf{v} = \mathbf{u}\left(\mathbf{v}^T\mathbf{v}\right), \tag{3.6}$$

which become

$$\left(\mathbf{u}^T\mathbf{v}\right)\mathbf{v} = 0,$$

since $\mathbf{v}$ is isotropic. Then, $\mathbf{v} \neq 0$ implies $\mathbf{u}^T\mathbf{v} = 0$. Noticing that $\mathrm{tr}(A) = \mathbf{u}^T\mathbf{v}$ finishes the proof. □

Moreover, we have a partial converse theorem.

Theorem 3.10 ([80], Theorem 2)**.** *Let A be a singular symmetric matrix with a null space of dimension 1. Then the null space contains an isotropic vector if the trace of its adjugate vanishes.*

Proof. As shown in the previous proof, since the dimension of the null space is 1 we can represent the adjoint of A as in (3.5), with $\mathbf{v} \neq 0$ a vector generating the null space and $\mathbf{u} \neq 0$. Using the symmetry of A equation (3.6) holds. Then, since $\mathbf{u}^T\mathbf{v} = \mathrm{tr}(A) = 0$ and $\mathbf{u} \neq 0$, $\mathbf{v}$ is an isotropic vector generating the null space of A. □

For completeness we state the converse of Theorem 3.9.

Theorem 3.11 ([80], Theorem 3)**.** *The null space of a singular symmetric matrix contains an isotropic vector if the trace of its adjugate vanishes.*

Proof. We can assume that the dimension of the null space is greater than 2, $\nu(A) \geq 2$. Indeed, the case $\nu(A) = 1$ has already been proved in Theorem 3.10. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ linearly independent vectors lying in the null space of A, then if one of them is isotropic we are done, otherwise we can normalized $\mathbf{v}_1$ onto $\hat{\mathbf{v}}_1$, such that $\hat{\mathbf{v}}_1^T\hat{\mathbf{v}}_1 = 1$. Therefore, we can use Gram-Schmidt orthogonalization procedure to obtain $\bar{\mathbf{v}}_2 = \mathbf{v}_2 - (\mathbf{v}_2^T\hat{\mathbf{v}}_1)\hat{\mathbf{v}}_1$. If $\bar{\mathbf{v}}_2$ is isotropic we are done. Otherwise, setting $\hat{\mathbf{v}}_2 = \bar{\mathbf{v}}_2/(\bar{\mathbf{v}}_2^T\bar{\mathbf{v}}_2)$ we get two vectors $\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2$ linearly independent, lying on the null space of A and such that $\hat{\mathbf{v}}_1^T\hat{\mathbf{v}}_2 = 0$ and $\hat{\mathbf{v}}_1^T\hat{\mathbf{v}}_1 = \hat{\mathbf{v}}_2^T\hat{\mathbf{v}}_2 = 1$. Then the vector $\hat{\mathbf{v}}_1 + i\hat{\mathbf{v}}_2$ is an isotropic vector in the null space of A. □

We will use the preceding theorems to give some properties of Jacobi matrices. First we state the following lemma.

Lemma 3.12. *Let λ be an eigenvalue of a complex Jacobi matrix J and $\mathbf{v}$ an associated eigenvector. Then, $\mathbf{v}$ is isotropic if and only if λ has an algebraic multiplicity greater than* 1.

Proof. Given a matrix $A(\xi)$ depending on a parameter ξ, *Jacobi's formula* states that

$$\frac{d}{d\xi}\det A(\xi) = \operatorname{tr}(\operatorname{adj}(A(\xi)\frac{dA(\xi)}{d\xi}));$$

for a proof see, e.g., [65, Theorem 1 at p. 149]. If $A(\xi) = \xi I - J$, then the previous formula becomes

$$\frac{d}{d\xi}\phi(\xi) = \operatorname{tr}(\operatorname{adj}(\xi I - J)),$$

where, ϕ is the characteristic polynomial of J. Let λ be an eigenvalue of J, then $(\lambda I - J)$ is a complex symmetric matrix such that

$$\det(\lambda I - J) = 0 \text{ and } \operatorname{tr}(\operatorname{adj}(\lambda I - J)) = \phi'(\lambda).$$

Since $\phi'(\lambda) = 0$ if and only if the algebraic multiplicity of λ is greater than 1, by Theorems 3.9 and 3.10 the eigenspace of J corresponding to λ contains an isotropic vector if and only if the algebraic multiplicity of λ is greater than 1. Since by Theorem 3.2 any complex Jacobi matrix is non-derogatory, the proof is finished. □

We summarize the situation in the following proposition.

Proposition 3.13. *If J is a Jacobi matrix, then the following properties are equivalent:*

1. *J is diagonalizable;*
2. *There exists a (complex) nonsingular matrix V such that $V^{-1} = V^T$ and $V^T J V$ is a diagonal matrix;*
3. *None of the eigenvectors of J is isotropic.*

Proof. The second and the third properties are equivalent by Theorem 3.8. Obviously the second one implies the first one. So it remains to prove that if J is diagonalizable, then no eigenvector is isotropic. Since J is non-derogatory, using Lemma 3.12 finishes the proof. □

Moreover, we can give a theorem for the non-diagonalizable case.

Proposition 3.14. *If J is a Jacobi matrix, then the following statements are equivalent:*

1. *J is not diagonalizable;*
2. *It does not exist a nonsingular matrix V such that $V^{-1} = V^T$ and $V^T J V$ is a Jordan form for J.*
3. *There exists an isotropic eigenvector of J.*

Proof. The statement 1. $\Leftrightarrow$ 3. has already been proved in Proposition 3.13. Furthermore, 2. $\Rightarrow$ 1. . In fact, 2. implies that it does not exist a nonsingular matrix V such that $V^{-1} = V^T$ and $V^T J V$ is a diagonal matrix. Thus, by Proposition 3.13, J is not diagonalizable.

We finish by proving that 3. $\Rightarrow$ 2. . Let Λ be the Jordan form of J and W such that $JW = W\Lambda$. Since J is non-derogatory, at least one column $\mathbf{w}_i$ of W is a multiple of the isotropic eigenvector. Since $\mathbf{w}_i^T \mathbf{w}_i = 0$ we see that $W^T W$ cannot be equal to the identity matrix. □

3.4 Moment Matching Property for Jacobi matrices

If a linear functional is defined by the moments $\mathcal{L}(x^i) = \mathbf{v}^* A^i \mathbf{v}, i = 0, 1, \ldots,$ where A is a Hermitian matrix, $\mathbf{v}$ is a nonzero vector and $\mathbf{v}^*$ is the conjugate transpose of $\mathbf{v}$. Then the following is a well-known result (e.g., refer to [42])

$$\mathcal{L}(x^i) = \mathbf{v}^* A^i \mathbf{v} = ||\mathbf{v}||^2 \, \mathbf{e}_1^T \, (J_n)^i \, \mathbf{e}_1, = m_0 \, \mathbf{e}_1^T \, (J_n)^i \, \mathbf{e}_1, \quad i = 0, 1, \ldots, 2n-1,$$

where J_n is the Jacobi matrix associated with the first n orthogonal polynomial with respect to $\mathcal{L}$ (for details refer to Chapter 1, in particular (1.10)). Using the Vorobyev method of moments [91, in particular Chapter III], this property can easily be extended, assuming the existence of the first n steps of the non-Hermitian Lanczos process (see Algorithm 5.4), to a general complex matrix A; see [85]. In this section we give a proof for an analogous property for Jacobi matrices determined by any quasi-definite linear functional.

Theorem 3.15. *[Moment Matching Property] Let $\mathcal{L}$ be a quasi-definite linear functional on $\mathcal{P}_n$ and let J_n be the Jacobi matrix of coefficients from the recurrence relations for orthogonal polynomials with respect to $\mathcal{L}$; see (1.9). Then*

$$\mathcal{L}(x^i) = m_0 \, \mathbf{e}_1^T (J_n)^i \, \mathbf{e}_1, \;\; i = 0, \ldots, 2n-1, \tag{3.7}$$

where $m_0 = \mathcal{L}(x^0)$.

We prove this theorem throughout the following two lemmas.

Lemma 3.16. *The polynomials $p_0, \dots, p_{n-1}$ associated with the three-term recurrence relation whose coefficients are given by the Jacobi matrix J_n are orthonormal with respect to the functional $\widetilde{\mathcal{L}}$ defined by*

$$\widetilde{\mathcal{L}}(x^i) = m_0\, \mathbf{e}_1^T\, (J_n)^i\, \mathbf{e}_1,$$

with $m_0 = 1/p_0^2$.

Proof. If J_n is a Jacobi matrix associated with the polynomials $p_0, \dots, p_{n-1}$, then for $i = 0, \dots, n-1$ the $(i+1)$-st entry of the vector $(J_n)^i \mathbf{e}_1$ is nonzero, and, for $i = 0, \dots, n-2$, the entries $i+2, \dots, n$ of $(J_n)^i \mathbf{e}_1$ are zero. Therefore the canonical basis $\mathbf{e}_1, \dots, \mathbf{e}_k$ is an orthonormal basis of the Krylov subspaces

$$\mathcal{K}_k(J_n, \mathbf{e}_1) = \operatorname{span}\{\mathbf{e}_1, J_n\mathbf{e}_1, \dots, (J_n)^{k-1}\, \mathbf{e}_1\}, \quad k = 1, \dots, n,$$

i.e., $\mathbf{e}_k = \tilde{p}_{k-1}(J_n)\, \mathbf{e}_1$ for a polynomial $\tilde{p}_{k-1}$ of degree $k-1$.

The polynomials $\hat{p}_{k-1} = \tilde{p}_{k-1}/\sqrt{m_0}$, $k = 1, \dots, n$, are orthonormal with respect to $\widetilde{\mathcal{L}}$. Indeed,

$$\widetilde{\mathcal{L}}(\hat{p}_i \hat{p}_j) = m_0\, \mathbf{e}_1^T \hat{p}_i(J_n) \hat{p}_j(J_n)\, \mathbf{e}_1 = \mathbf{e}_i^T \mathbf{e}_j.$$

From $\mathbf{e}_1 = \tilde{p}_0(J_n)\, \mathbf{e}_1$ we obtain $\tilde{p}_0 \equiv 1$, or equivalently, $\hat{p}_0 = 1/\sqrt{m_0} = p_0$. We finally show that $\hat{p}_k = p_k$ for $k = 1, \dots, n-1$. Notice that

$$\widetilde{\mathcal{L}}(x\hat{p}_i \hat{p}_j) = m_0\, \mathbf{e}_1^T \hat{p}_i(J_n) J_n \hat{p}_j(J_n)\, \mathbf{e}_1 = (J_n)_{i,j}.$$

Hence, by (1.9) the coefficients from the three-term recurrence relation for $x\hat{p}_0, \dots, x\hat{p}_{n-1}$ are the same as those for $xp_0, \dots, xp_{n-1}$. And the proof is finished. □

The following lemma finishes the proof of Theorem 3.15.

Lemma 3.17. *Let $\mathcal{L}$ and $\widetilde{\mathcal{L}}$ be linear functionals such that there exists a sequence of polynomials p_i for $i = 0, \dots, n-1$ that are orthogonal with respect to both $\mathcal{L}$ and $\widetilde{\mathcal{L}}$. Let $\mathcal{L}(x^0) = \widetilde{\mathcal{L}}(x^0)$. Then*

$$\mathcal{L}(x^i) = \widetilde{\mathcal{L}}(x^i) \text{ for } i = 0, \dots, 2n-1. \tag{3.8}$$

Proof. We prove the result by induction. Using (1.13),

$$\frac{\mathcal{L}(xp_0^2(x))}{\mathcal{L}(p_0^2(x))} = \alpha_0 = \frac{\widetilde{\mathcal{L}}(xp_0^2(x))}{\widetilde{\mathcal{L}}(p_0^2(x))}, \quad \text{i.e.,} \quad \frac{m_1}{m_0} = \frac{\widetilde{m}_1}{\widetilde{m}_0}.$$

Since we have assumed $m_0 = \widetilde{m}_0$, we conclude $m_1 = \widetilde{m}_1$. Let $m_i = \widetilde{m}_i$ for $i = 0, \ldots, 2k-3$. Using (1.13) we have

$$\frac{\mathcal{L}(xp_{k-1}(x)p_{k-2}(x))}{\mathcal{L}(p_{k-2}^2(x))} = \gamma_{k-1} = \frac{\widetilde{\mathcal{L}}(xp_{k-1}(x)p_{k-2}(x))}{\widetilde{\mathcal{L}}(p_{k-2}^2(x))}.$$

Rewriting

$$xp_{k-1}(x)p_{k-2}(x) = \sum_{i=0}^{2k-2} a_i x^i \quad \text{and} \quad p_{k-2}^2(x) = \sum_{i=0}^{2k-4} b_i x^i,$$

the induction assumptions give $m_{2k-2} = \widetilde{m}_{2k-2}$. Repeating the same argument with the coefficient α_{k-1} finishes the proof. □

A different normalization of the orthogonal polynomials is associated with a tridiagonal matrix T_n such that $J_n = D^{-1}T_nD$, with D the diagonal matrix with elements given in (1.22) (see also Proposition 1.6). Hence, the statement of Theorem 3.15 remains valid for any tridiagonal matrix T_n associated with a sequence of orthogonal polynomials defined by the functional $\mathcal{L}$.

3.5 Real Jacobi Matrices

When we deal with real Jacobi matrices we have some important additional properties, which we summarize in this section. First of all we notice that any real $n \times n$ Jacobi matrix J_n is a symmetric matrix. Thus it has real eigenvalues and can be orthogonally diagonalized, i.e.,

$$J_n W = W \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of the matrix J_n and $W = [\mathbf{w}_1, \ldots, \mathbf{w}_n] \in \mathbb{R}^{n\times n}$ is an orthogonal matrix whose columns are the normalized eigenvectors of J_n, $W^TW = WW^T = I$. Then, Proposition 3.5 and Corollary 3.3 give the following statement.

Theorem 3.18. *The following properties hold for every real Jacobi matrix:*

1. *Eigenvalues are real and distinct;*
2. *The first and the last component of each of its eigenvectors are nonzero.*

As we have seen in Chapter 1, every Jacobi matrix is associated to a sequence of orthonormal polynomials $p_0, \ldots, p_n$ (see Favard Theorem 1.5).

Moreover, by (1.9) the eigenvalues J_n are the roots of p_n. Hence, the *Strict Interlacing Property 1.16* holds for the eigenvalues of any real Jacobi matrix.

Moreover, let $J_1, \ldots, J_n$ be Jacobi matrices such that J_i is the leading principal $i \times i$ submatrix of J_n for $i = 1, \ldots, n-1$. From now on $J_1, \ldots, J_n$ will always denote the described sequence of Jacobi matrices.

Theorem 3.19 (Interlacing Property)**.** *Let $J_1, \ldots, J_n$ be real Jacobi matrices as described above. Let k and ℓ be integers with $k+1 \leq \ell \leq n$ and let $\lambda_i^{(k)}$ for $i = 1, \ldots, k$ be the eigenvalues of J_k. Then at least one of the eigenvalues of J_ℓ is contained in any of the $k+1$ open intervals*

$$\left(-\infty, \lambda_1^{(k)}\right), \left(\lambda_1^{(k)}, \lambda_2^{(k)}\right), \ldots, \left(\lambda_{k-1}^{(k)}, \lambda_k^{(k)}\right), \left(\lambda_k^{(k)}, +\infty\right).$$

The proof of this theorem uses Gauss quadrature and can be found in [63, Theorem 3.3.1, p. 92 and Remark 3.4.4, p. 115]. It is important to remark that the proof cannot be extended to non positive definite linear functionals. As a trivial consequence we get the strict interlacing property for the eigenvalues of two subsequent Jacobi matrices J_k and J_{k+1} and, equivalently, the strict interlacing property of the roots of two consecutive orthogonal polynomials.

CHAPTER 4

Gauss Quadrature for Linear Functionals

4.1 Gauss Quadrature under Restrictive Assumptions

Let $\mathcal{L}$ be a *Positive definite* linear functional (see Definition 1.10). Let f be a function from the space on which $\mathcal{L}$ is defined, then we can approximate the value $\mathcal{L}(f)$ with the n-node quadrature rule

$$\mathcal{L}(f) \approx \sum_{i=1}^{n} \omega_i f(\lambda_i),$$

where $\lambda_1, \ldots, \lambda_n$ are the nodes and $\omega_1, \ldots, \omega_n$ are the weights. With a particular choice of the nodes and the weights, depending only on $\mathcal{L}$, the quadrature rule is exact for every polynomial f of degree lower than or equal to $2n-1$. In this case we have a *Gauss quadrature rule*. The classical theory of Gauss quadrature can be found in many books; see, for example, [87, Chapters III and XV], [15, Chapter I, Section 6], [35], [36, Chapter 3.2], [63, Section 3.2].

We recall that in the classical case (see [33], [54], [16], [17] and [84]) $\mathcal{L}$ is the Riemann, the weighted Riemann or the more general Riemann-Stieltjes integral with respect to a non-decreasing distribution function μ defined on the real axis having finite limits at $\pm\infty$ and infinitely many points of increase. Since μ is of bounded variation, the integral $\int f \mathrm{d}\mu$ exists for every continuous function f. However, every positive definite linear functional can be seen as an integral with respect to a positive non-decreasing distribution function supported on the real axis; see Appendix A and Section 1.2.

Let us recall some basic properties of *Gauss quadrature*:

- G1: The n-node Gauss quadrature rule attains the maximal algebraic degree of exactness $2n-1$.
- G2: The n-node Gauss quadrature is well-defined and it is unique. Naturally, the Gauss quadrature rules with a smaller number of nodes also exist and they are unique.
- G3: The Gauss quadrature of the function f can be written in the form $m_0\,\mathbf{e}_1^T f(J_n)\mathbf{e}_1$, where J_n is the Jacobi matrix containing the coefficients from the three-term recurrence relation for *orthonormal* polynomials associated with $\mathcal{L}$; $m_0 = \mathcal{L}(x^0)$.

Since the degree of exactness is larger than $n-1$, the Gauss quadrature is an interpolatory quadrature, i.e., the weights ω_i satisfy

$$\omega_i = \mathcal{L}(\ell_i),\ i = 1, \ldots, n, \tag{4.1}$$

where $\ell_i(x)$ is the *Lagrange interpolation polynomial*, defined as

$$\ell_i(x) = \frac{(x-\lambda_1)\ldots(x-\lambda_{i-1})(x-\lambda_{i+1})\ldots(x-\lambda_n)}{(\lambda_i-\lambda_1)\ldots(\lambda_i-\lambda_{i-1})(\lambda_i-\lambda_{i+1})\ldots(\lambda_i-\lambda_n)} = \frac{\pi_n(x)}{(x-\lambda_i)\pi_n'(\lambda_i)}.$$

In fact, since $\ell_i(x)$ has degree $n-1$, then

$$\mathcal{L}(\ell_i) = \sum_{j=1}^{n} \omega_j \ell_i(\lambda_j) = \omega_i,$$

for $i = 1, \ldots, n$.

We will now revisit the situation for the functional $\mathcal{L}$ that is only quasi-definite. We start with the usual form of an n-node quadrature rule

$$\mathcal{L}(f) = \sum_{i=1}^{n} \omega_i f(\lambda_i) + R_n(f), \tag{4.2}$$

where the nodes $\lambda_1, \ldots, \lambda_n$ are distinct and the last term stands for the quadrature error.

Theorem 4.1. *The quadrature rule (4.2) is exact for every f in $\mathcal{P}_{2n-1}$ if and only if it is interpolatory and the polynomial*

$$\varphi_n(x) = \prod_{i=1}^{n}(x-\lambda_i) \tag{4.3}$$

satisfies $\mathcal{L}(\varphi_n p) = 0$ for every $p \in \mathcal{P}_{n-1}$.

Proof. Let (4.2) be exact for every $f \in \mathcal{P}_{2n-1}$. Then for every $p \in \mathcal{P}_{n-1}$ we get $R_n(\varphi_n p) = 0$ and moreover, since $\varphi_n(\lambda_i) = 0$ for $i = 1, ..., n$,

$$\mathcal{L}(\varphi_n p) = \sum_{i=1}^{n} \omega_i \varphi_n(\lambda_i) p(\lambda_i) = 0.$$

Vice versa, let $\mathcal{L}(\varphi_n p) = 0$ for all p from $\mathcal{P}_{n-1}$. Since any $f \in \mathcal{P}_{2n-1}$ can be written as

$$f(x) = \varphi_n(x) q(x) + r(x) \tag{4.4}$$

for some q and r from $\mathcal{P}_{n-1}$, $\mathcal{L}(f) = \mathcal{L}(r)$. The quadrature is an interpolatory quadrature on n nodes, hence it has algebraic degree of exactness at least $n-1$. Thus $\mathcal{L}(r) = \sum_{i=1}^{n} \omega_i \, r(\lambda_i)$. Since $\varphi_n(\lambda_i) = 0$, by (4.4) we get $r(\lambda_i) = f(\lambda_i)$ for $i = 1, \ldots, n$. This finishes the proof. □

We can interpret Theorem 4.1 with the orthogonal polynomials theory presented in Chapter 1. The monic polynomial φ_n of degree n has n distinct roots $\lambda_1, \ldots, \lambda_n$. Moreover, it is orthogonal to the space $\mathcal{P}_{n-1}$ with respect to the linear functional $\mathcal{L}$, i.e., $\mathcal{L}(\varphi_n p) = 0$ for every p from $\mathcal{P}_{n-1}$. Then the interpolatory quadrature with the nodes $\lambda_1, \ldots, \lambda_n$ has algebraic degree of exactness at least $2n-1$. We recall that a sequence of $n+1$ orthogonal polynomials exists if and only if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$; see Theorem 1.3. Therefore, the quadrature rule (4.2) has the properties G1 and G2 if and only if the following conditions simultaneously hold:

1. There exists a sequence of orthogonal polynomials $p_0, \ldots, p_n$ with respect to the linear functional $\mathcal{L}$ (i.e., $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$);

2. Zeros of the individual polynomials p_j, $j = 1, \ldots, n$, in the sequence are distinct;

Let $J_1, \ldots, J_n$ be the Jacobi matrices associated with a sequence of orthonormal polynomials with respect to $\mathcal{L}$; see (1.10). Then by Corollary 3.3 and Theorem 1.3 the conditions 1. and 2. are respectively equivalent to the following ones.

1. $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$;

2. The Jacobi matrices $J_1, \ldots, J_n$ associated with $\mathcal{L}$ are diagonalizable.

Assuming that $\mathcal{L}$ satisfy conditions 1. and 2., by the moment matching Theorem 3.15 we know that for every polynomial $f \in \mathcal{P}_{2n-1}$

$$\mathcal{L}(f) = m_0 \, \mathbf{e}_1^T f(J_n) \, \mathbf{e}_1,$$

with $m_0 = \mathcal{L}(x^0)$. Since we assume J_n diagonalizable, then $\Lambda = V^{-1} J_n V$ is diagonal. Moreover, Proposition 3.13 implies $V^{-1} = V^T$. By the definition of matrix function (Definition 2.2)

$$\mathbf{e}_1^T V \begin{bmatrix} f(\lambda_1) & 0 & \cdots & 0 \\ 0 & f(\lambda_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & f(\lambda_n) \end{bmatrix} V^T \mathbf{e}_1 = \sum_{i=1}^{n} (v_i)^2 f(\lambda_i), \tag{4.5}$$

where v_i is the first element of the i-th column of V, for $i = 1, \ldots, n$. By Property G2 the n-node quadrature rule (4.2) is unique. Hence, the quadrature (4.2) can be expressed in the form $m_0\, \mathbf{e}_1^T f(J_n)\, \mathbf{e}_1$, i.e., the property G3 is generalized in a straightforward way; see [35, p. 153], [78, p. 267-268]. Moreover, using (4.5) and the uniqueness of the quadrature rule (4.2) we get

$$\omega_i = m_0 (v_i)^2, \qquad \text{for } i = 1, \ldots, n.$$

Hence, the weights of the quadrature rule are m_0 times the square of the first element of the eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of the Jacobi matrix J_n, normalized such that $\mathbf{v}_i^T \mathbf{v}_i = 1$. Of course, the nodes are the eigenvalues of J_n.

When $\mathcal{L}$ is positive definite, J_n and V are real matrices. Hence, we recover the well-known property $\omega_i = m_0 (v_i)^2 > 0$, for every weight ω_i of a Gauss quadrature rule; see, e.g., [94, Sections 2.5 and 2.9] and [43].

To our knowledge quadrature (4.2) was considered for the first time by Gragg in [44] for real valued linear functionals. A generalization for complex valued functionals was considered by Saylor and Smolarski in [78]. However, due to the assumption on the distinctness of the nodes, see Property 2, this construction is restrictive. Indeed, if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_k$, then the orthogonal polynomials in the sequence $p_1, \ldots, p_k$ *can have multiple zeros*. Hence it can happen that for some values ℓ, $\ell \leq k$, the ℓ-point quadrature defined by

$$\mathcal{L}(f) \approx \sum_{i=1}^{\ell} \omega_i f(\lambda_i)$$

cannot be properly defined, i.e., it represents an interpolatory quadrature on strictly less than ℓ distinct points. Thus it cannot achieve the algebraic degree of exactness $2\ell - 1$. We demonstrate this with the following example.

Example 4.1 Let $\mathcal{L}$ a linear functional defined by a sequence of moments with the first seven terms given by

$$1,\ 3,\ 8,\ 20,\ 52,\ 156,\ i,$$

as in Example 3.1. Then $\mathcal{L}$ is quasi-definite on $\mathcal{P}_3$, since

$$\Delta_0 = 1, \quad \Delta_1 = -1, \quad \Delta_2 = -4, \quad \Delta_3 = 2128 - 4i.$$

The associated monic orthogonal polynomials are

$$\pi_0 = 1, \quad \pi_1(x) = x - 3, \quad \pi_2(x) = x^2 - 4x + 4, \quad \pi_3(x) = x^3 - 7x^2 + 20x - 24.$$

The zeros of π_2 are $\lambda_1 = \lambda_2 = 2$, which means that the 2-node quadrature (4.2) which is exact on $\mathcal{P}_3$ does not exist. However, the zeros of π_3 are $\lambda_1 = 3$, $\lambda_2 = 2 - 2i$ and $\lambda_3 = 2 + 2i$, which means that there exists the 3-node quadrature (4.2) which is exact on $\mathcal{P}_5$. The corresponding Jacobi matrix is

$$J_3 = \begin{bmatrix} 3 & i & 0 \\ i & 1 & 2i \\ 0 & 2i & 3 \end{bmatrix}.$$

The matrix J_3 is diagonalizable, whereas its leading principal 2×2 submatrix is not.

4.2 n-weight Gauss Quadrature

To overcome restrictions for quasi-definite linear functionals that produce diagonalizable Jacobi matrices, we need to modify the quadrature concept of relation (4.2). We then consider the *n-weight quadrature formula*

$$\mathcal{L}(f) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i - 1} \omega_{i,j} \, f^{(j)}(\lambda_i) + R_n(f), \tag{4.6}$$

with $n = s_1 + \ldots + s_\ell$. We remark that quadrature (4.2) is the special case of the quadrature (4.6) when $\ell = n$ and $s_1 = \ldots = s_n = 1$. So we are generalizing the rule (4.2) in the way that considers, in addition to the function values $f(\lambda_1), \ldots, f(\lambda_\ell)$, also the values of the derivatives of f at the points $\lambda_1, \ldots, \lambda_\ell$. Therefore the quadrature (4.6) needs more smoothness of the argument function f in $\mathcal{L}(f)$. We explain the construction (4.6) using the following theorems that show how to choose the values $s_1, \ldots, s_\ell$ when we want to achieve the maximal degree of exactness.

Theorem 4.2 ([74]). *Let $\mathcal{L}$ be an arbitrary linear functional on $\mathcal{P}$. The quadrature rule (4.6) is exact for every f in $\mathcal{P}_{2n-1}$ if and only if it is exact for $\mathcal{P}_{n-1}$ and the polynomial*

$$\varphi_n(x) = (x - \lambda_1)^{s_1} (x - \lambda_2)^{s_2} \ldots (x - \lambda_\ell)^{s_\ell} \tag{4.7}$$

satisfies $\mathcal{L}(\varphi_n \, p) = 0$ for every $p \in \mathcal{P}_{n-1}$.

Proof. As done in [83] for each root $\lambda_1, \ldots, \lambda_\ell$ of φ_n we define the polynomials $h_{i,j}$ of degree $n-1$

$$h_{i,j}(x) = \frac{(x-\lambda_i)^j}{j!} \left\{ \sum_{\nu=0}^{s_i-1-j} \frac{(x-\lambda_i)^\nu}{\nu!} \left(\frac{1}{g_i(x)} \right)^{(\nu)} \Bigg|_{x=\lambda_i} \right\} g_i(x), \qquad (4.8)$$
$$j = 0, 1, \ldots, s_i - 1,$$

with $g_i(x) = \prod_{\substack{t=1 \\ t \neq i}}^{\ell} (x-\lambda_t)^{s_t}$. From (4.8) we obtain

$$h_{i,j}^{(t)}(\lambda_k) = 1 \quad \text{for } \lambda_k = \lambda_i \text{ and } t = j,$$
$$h_{i,j}^{(t)}(\lambda_k) = 0 \quad \text{for } \lambda_k \neq \lambda_i \text{ or } t \neq j,$$

where $k = 1, 2, \ldots, \ell$, and $t = 0, 1, \ldots, s_i - 1$; see [82, Section 3]. Now, we can define the generalized (Hermite) interpolating polynomial as

$$h_{n-1}(x) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} f^{(j)}(\lambda_i) h_{i,j}(x);$$

we refer to [82]. So we get that the formula (4.6) is exact for any polynomial $f \in \mathcal{P}_{n-1}$ if and only if

$$\mathcal{L}(f) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} w_{i,j} f^{(j)}(\lambda_i) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \mathcal{L}(h_{i,j}) f^{(j)}(\lambda_i),$$

i.e., if and only if the weights of the quadrature (4.6) can be expressed as

$$\omega_{i,j} = \mathcal{L}(h_{i,j}).$$

The remaining part of the proof is completely similar to the proof of Theorem 4.1. □

We say that the n-weight quadrature rule (4.6) is unique if the nodes and the weights are uniquely determined by $\mathcal{L}$ and n.

Theorem 4.3 ([74])**.** *Let $\mathcal{L}$ be an arbitrary linear functional on $\mathcal{P}$. The n-weight quadrature (4.6) of degree of exactness at least $2n-1$ exists and is unique if and only if the n-th Hankel determinant (1.4) is nonzero, i.e., $\Delta_{n-1} \neq 0$.*

Proof. By Theorem 4.2 the n-weight interpolatory quadrature (4.6) is of degree of exactness at least $2n-1$ if and only if the monic polynomial (4.7)

$$\varphi_n(x) = x^n + c_{n-1}x^{n-1} + \ldots + c_1 x + c_0$$

is orthogonal to the space $\mathcal{P}_{n-1}$. The conditions $\mathcal{L}(x^j\varphi_n) = 0$, $j = 0, \ldots, n-1$ then are equivalent to the system

$$\begin{bmatrix} m_0 & m_1 & \ldots & m_{n-1} \\ m_1 & m_2 & \ldots & m_n \\ \vdots & \vdots & \ddots & \vdots \\ m_{n-1} & m_n & \ldots & m_{2n-2} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} -m_n \\ -m_{n+1} \\ \vdots \\ -m_{2n-1} \end{bmatrix}, \tag{4.9}$$

which has a unique solution if and if and only if $\Delta_{n-1} \neq 0$. □

Finally, we give the condition under which the degree of exactness of (4.6) is exactly $2n-1$ (i.e., it does not exceed $2n-1$). This has no counterpart in the positive-definite case, in which the n-node Gauss quadrature cannot have algebraic degree of exactness larger than $2n-1$.

Theorem 4.4 ([74])**.** *Let $\mathcal{L}$ be an arbitrary linear functional on $\mathcal{P}$ and let the n-weight quadrature (4.6) has degree of exactness at least $2n-1$. Then the degree of exactness of the quadrature (4.6) is (exactly) $2n-1$ if and only if the $(n+1)$-st Hankel determinant (1.4) is nonzero, i.e., $\Delta_n \neq 0$.*

Proof. The n-weight quadrature rule (4.6) has degree of exactness at least $2n-1$. Then the polynomial φ_n (4.7) is orthogonal to $\mathcal{P}_{n-1}$. In addition, φ_n is orthogonal to $\mathcal{P}_n$ if and only if $\mathcal{L}(\varphi_n^2) = 0$ in which case the degree of exactness of (4.6) is at least $2n$. Therefore, the quadrature (4.6) has degree of exactness larger than $2n-1$ if and only if $\mathcal{L}(\varphi_n x^j) = 0$ for $j = 0, \ldots, n$, or equivalently, if and only if there is a vector $[c_0, \ldots, c_{n-1}, 1]^T$ such that

$$\begin{bmatrix} m_0 & m_1 & \ldots & m_n \\ m_1 & m_2 & \ldots & m_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_n & m_{n+1} & \ldots & m_{2n} \end{bmatrix} \begin{bmatrix} c_0 \\ \vdots \\ c_{n-1} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \tag{4.10}$$

which implies $\Delta_n = 0$. □

Corollary 4.5 ([74])**.** *The quadrature rule (4.6) has the properties G1 and G2 if and only if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$.*

Proof. The n-weight quadrature rule (4.6) is unique and of degree of exactness $2n-1$ if and only if both Δ_{n-1} and Δ_n are different from zero. The property G2 requires the same for all j-weight quadratures with $j = 1, \ldots, n-1$, and thus all Hankel determinants Δ_j, $j = 0, \ldots, n$ have to be nonzero and this is equivalent to ask $\mathcal{L}$ to be quasi-definite on $\mathcal{P}_n$. □

Now we want to show that property G3 is true for the quadrature rule (4.6) when $\mathcal{L}$ is quasi definite. Let J_n be an $n \times n$ Jacobi matrix, with λ_i its eigenvalues of algebraic multiplicities s_i, $i = 1, \ldots, \ell$. By Theorem 3.2 the matrix J_n is non-derogatory. Hence, its Jordan normal form (2.1) $W^{-1} J_n W = [\Lambda_1, \ldots, \Lambda_\ell]$, has only one Jordan block Λ_i for every distinct eigenvalue λ_i, for $i = 1, \ldots, \ell$. Recalling the definition of a matrix function (Definition 2.2), and denoting the first row of W as

$$\mathbf{w}^T = [w_{1,0}, \ldots, w_{1,s_1-1}, w_{2,0}, \ldots, w_{2,s_2-1}, \ldots, w_{\ell,0}, \ldots, w_{\ell,s_\ell-1}],$$

and the first column of W^{-1} as

$$\hat{\mathbf{w}} = [\hat{w}_{1,0}, \ldots, \hat{w}_{1,s_1-1}, \hat{w}_{2,0}, \ldots, \hat{w}_{2,s_2-1}, \ldots, \hat{w}_{\ell,0}, \ldots, \hat{w}_{\ell,s_\ell-1}]^T,$$

we obtain

$$\begin{aligned} \mathbf{e}_1^T f(J_n)\, \mathbf{e}_1 &= \mathbf{e}_1^T\, W \mathrm{diag}(f(\Lambda_1), \ldots, f(\Lambda_\ell))\, W^{-1} \mathbf{e}_1 \\ &= \mathbf{w}^T \mathrm{diag}(f(\Lambda_1), \ldots, f(\Lambda_\ell))\, \hat{\mathbf{w}} \\ &= \sum_{i=1}^{\ell} [w_{i,0}, \ldots, w_{i,s_i-1}] f(\Lambda_i) [\hat{w}_{i,0}, \ldots, \hat{w}_{i,s_i-1}]^T. \end{aligned}$$

Equation (3.3) in Chapter 3 gives an explicit form for the eigenvectors of tridiagonal matrices with nonzero elements on the sub- and super-diagonal. Similarly Proposition 3.4 gives an explicit form for generalized eigenvectors of the same kind of matrices. From these results we deduce that the first elements of the columns of the matrix W are zero except for the columns that are eigenvectors of J_n. Moreover we remark that the first element of every eigenvector of such a matrix is nonzero (see Proposition 3.5). Therefore, the individual terms in the previous sum can be written as

$$[w_{i,0}, 0, \ldots, 0] \begin{bmatrix} f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \ldots & \frac{f^{(s_i-1)}(\lambda_i)}{(s_i-1)!} \\ 0 & f(\lambda_i) & \ldots & \frac{f^{(s_i-2)}(\lambda_i)}{(s_i-2)!} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & f(\lambda_i) \end{bmatrix} \begin{bmatrix} \hat{w}_{i,0} \\ \hat{w}_{i,1} \\ \vdots \\ \hat{w}_{i,s_i-1} \end{bmatrix}.$$

Hence we have

$$\mathbf{e}_1^T f(J_n)\,\mathbf{e}_1 = \sum_{i=1}^{\ell}\sum_{j=0}^{s_i-1} \frac{w_{i,0}\hat{w}_{i,j}}{j!}\, f^{(j)}(\lambda_i) = \sum_{i=1}^{\ell}\sum_{j=0}^{s_i-1} \tilde{\omega}_{i,j}\, f^{(j)}(\lambda_i), \tag{4.11}$$

with

$$\tilde{\omega}_{i,j} = \frac{w_{i,0}\hat{w}_{i,j}}{j!}, \quad \text{for } i = 1,\ldots,\ell, \quad j = 0,\ldots,s_i-1.$$

Using $\omega_{i,j} = m_0\tilde{\omega}_{i,j}$ in (4.11) we get

$$m_0\mathbf{e}_1^T f(J_n)\,\mathbf{e}_1 = \sum_{i=1}^{\ell}\sum_{j=0}^{s_i-1} \omega_{i,j}\, f^{(j)}(\lambda_i). \tag{4.12}$$

We can prove the following corollary.

Corollary 4.6 ([74])**.** *The quadrature rule (4.6) having the properties G1 and G2 satisfies also the property G3.*

Proof. Notice that the right-hand side of (4.12) is of the form (4.6). Hence, it is enough to prove that the weights $\omega_{i,j}$ are equal to $\mathcal{L}(h_{i,j})$, with polynomials $h_{i,j}$ defined by (4.8) (see the proof of Theorem 4.2). Quadrature (4.6) satisfies the properties G1 and G2, thus, using Corollary 4.5, the functional $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$. Moreover, by Theorem 3.15 its values on monomials x^i must then be equal for $i = 0, 1, \ldots, 2n-1$ to the right-hand side of (4.12) with $f(\lambda)$ replaced by the same monomials. Therefore, the right-hand side of (4.12) is a quadrature rule of algebraic degree at least $2n-1$. Using uniqueness it must be equal to the quadrature rule (4.6) with the weights $\mathcal{L}(h_{i,j})$. □

The following theorem summarizes the results about the relation between the n-weight quadrature formula (4.6) and the associated Jacobi matrix.

Theorem 4.7 ([74])**.** *Let $\mathcal{L}$ be an arbitrary linear functional on $\mathcal{P}$ and $m_0 = \mathcal{L}(x^0)$. There exists a Jacobi matrix J_n of dimension n such that*

$$\mathcal{L}(x^i) = m_0\,\mathbf{e}_1^T (J_n)^i\,\mathbf{e}_1, \quad \textit{for } i = 0,\ldots,2n-1, \tag{4.13}$$

$$\mathcal{L}(x^{2n}) \neq m_0\,\mathbf{e}_1^T (J_n)^{2n}\,\mathbf{e}_1, \tag{4.14}$$

if and only if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$.

Proof. If J_n is the Jacobi matrix satisfying (4.13) and (4.14), then, by (4.12) there exists the n-weight quadrature rule (4.6) with the degree of exactness $2n-1$. Therefore, Theorem 4.4 implies $\Delta_n \neq 0$. We need to prove that $\mathcal{L}$ is quasi-definite on $\mathcal{P}_{n-1}$ in order to show that $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$. Let

$p_0, \ldots, p_{n-1}$ be the polynomials associated with the three-term recurrence relation whose coefficients are given by J_n. Then, using Lemma 3.16 they are orthonormal with respect to the linear functional

$$\tilde{\mathcal{L}}(f) = m_0 \, \mathbf{e}_1^T f(J_n) \, \mathbf{e}_1, \quad \text{for } f \in \mathcal{P}.$$

In addition, they are orthonormal with respect to $\mathcal{L}$ by (4.13), i.e., $\mathcal{L}$ is quasi-definite on $\mathcal{P}_{n-1}$, by Theorem 1.3. The converse statement directly follows from corollaries 4.5 and 4.6. □

If the linear functional is quasi-definite, construction (4.6) and the related statements proved in this section show that it is possible to construct the n-weight quadrature (4.6) satisfying the properties G1,G2 and G3 of the classical Gauss quadrature. We remark that the quadrature (4.6) is different from the Gauss quadrature with multiple nodes considered in [14] and [73], and later in [41]. In particular, the latter assumes *positive-definite linear functionals* and has degree of exactness equal to

$$(\text{the number of weights}) + (\text{the number of nodes}) - 1.$$

The Gauss quadrature proposed in this section is constructed for *quasi-definite linear functionals* and has the degree of exactness

$$2 \times (\text{the number of weights}) - 1$$

that is larger than in the previous case.

We have proved that quasi-definiteness of $\mathcal{L}$ is a necessary and sufficient condition for the n-weight quadrature rule (4.6) to have all three properties G1, G2 and G3. Thus, for *non-definite linear functionals* all three properties cannot be satisfied.

Let $\mathcal{L}$ be a linear functional for which the n-th Hankel determinant (1.4) is equal to zero, i.e., $\Delta_{n-1} = 0$. By Theorem 4.3, the n-weight quadrature (4.6) having degree of exactness at least $2n-1$ either does not exist (the system (4.9) has no solution), or there are infinitely many of them (the system (4.9) has infinitely many solutions). Therefore the property G2 cannot be satisfied. If there exist infinitely many n-weight quadrature rules (4.6), then Δ_n must be equal to zero. Indeed, by (4.9), the first n rows of the matrix of the system (4.10) are linearly dependent. Thus, by Theorem 4.4 the degree of exactness of the n-weight quadratures (4.6) is greater than or equal to $2n$. Hence the property G1 is not satisfied as well.

In addition, if n is the smallest index for which $\Delta_{n-1} = 0$, then there exists a unique $(n-1)$-weight quadrature Q_{n-1} of the form (4.6) having degree of exactness at least $2n-3$. However, by Theorem 4.4 it does not

satisfy the property G1 since its degree of exactness is larger than $2n-3$. In the quasi-definite case the degree of exactness is uniquely determined; see theorems 4.3 and 4.4. While, with the $(n-1)$-weight quadrature Q_{n-1} the situation is different. If we only know the moments $m_0, \ldots, m_{2n-2}$, then it is not possible to determine the degree of exactness of Q_{n-1}. Indeed, if $Q_{n-1}(x^{2n-1}) \neq m_{2n-1}$, then the degree of exactness is $2n-2$. However, if $Q_{n-1}(x^{2n-1}) = m_{2n-1}$, then the degree of exactness of Q_{n-1} is at least $2n-1$, and so on. The following example show this fact.

Example 4.2 Let the linear functional $\mathcal{L}$ from Example 4.1 be defined by a sequence of moments with the first seven terms given by

$$1,\ 3,\ 8,\ 20,\ 52,\ 156,\ i.$$

Then, $\mathcal{L}$ is quasi-definite on $\mathcal{P}_3$. Moreover, we saw above that the 2-node quadrature (4.2) of degree of exactness 3 does not exist, since the zeros of π_2 are $x_1 = x_2 = 2$. However, we can consider the 2-weight quadrature rule of the form (4.6), i.e., $\omega_1 f(2) + \omega_2 f'(2)$. Since $\Delta_1 \neq 0$, by Theorem 4.3 the nonlinear system $\omega_1 z^j + j\omega_2 z^{j-1} = m_j$ for monomials $1, z, z^2$ and z^3, i.e.,

$$\begin{aligned} \omega_1 \cdot 1 + \omega_2 \cdot 0 &= 1 \\ \omega_1 z + \omega_2 \cdot 1 &= 3 \\ \omega_1 z^2 + 2\omega_2(z) &= 8 \\ \omega_1 z^3 + 3\omega_2(z^2) &= 20 \end{aligned}$$

has a unique solution (in $\mathbb{C}$): $\omega_1 = 1, \omega_2 = 1, z_1 = 2$. Moreover, since $\Delta_2 \neq 0$ by Theorem 4.4 the quadrature $f(2) + f'(2)$ has degree of exactness 3. We would have an higher degree of exactness if and only if $m_4 = 2^4 + 4 \cdot 2^3 = 48$. However, we would have $\Delta_2 = 0$, i.e., $\mathcal{L}$ would not be quasi-definite on $\mathcal{P}_2$. Furthermore, if $m_5 = 2^5 + 5 \cdot 2^4 = 112$, then the quadrature $f(2) + f'(2)$ would have degree of exactness at least 5, and so on.

The goal of this part of the thesis is to see how far we can go with generalization of the Gauss quadrature as an approximant for arbitrary linear functionals. In order to define some minimal properties of Gauss quadrature we proposed that any (generalization of the) Gauss quadrature should have the properties G1–G3. Hence, in this sense, we showed that the quasi-definiteness of the linear functional represents the *necessary and sufficient condition* for the existence of the Gauss quadrature. Hence, we will call an *n-weight Gauss Quadrature* the quadrature rule (4.6), and this is the quadrature for linear functionals quasi-definite on $\mathcal{P}_n$ which gives the maximal possible extension of this concept.

CHAPTER 5

Lanczos Algorithms

The goal of this chapter is to approximate the bilinear form

$$\mathbf{w}^* f(A)\mathbf{v}, \tag{5.1}$$

with A a complex matrix, $\mathbf{w}, \mathbf{v}$ complex vectors and f a matrix function (see Chapter 2). In particular, we use the results presented in the previous chapters to show how Lanczos algorithms can compute an approximation of (5.1). Lower and upper bounds for (5.1) are well-known when A is an Hermitian matrix and $\mathbf{w} = \mathbf{v} \neq 0$; we refer to [42, Chapter 7]. It is possible to extend the approximation to the non-Hermitian case, i.e., when A is not Hermitian and $\mathbf{w}$ and $\mathbf{v}$ can be different, using the non-Hermitian Lanczos algorithm (this was proved throughout the Vorobyev moment method in [85]). Furthermore, we want to show the strict relationship between the approximation of (5.1) in the non-Hermitian case and the n-weight Gauss quadrature rule introduced in Chapter 4, see in particular (4.6).

In Section 5.1 we recall that the space $\mathcal{P}_{n-1}$ of polynomial of degree at most $n-1$ is isomorphic to Krylov subspaces of dimension n, under some assumptions. Then we obtain Lanczos algorithms through the Stieltjes procedure for the computation of orthogonal polynomials (Section 5.2). In Section 5.3 we show the connection between moment matching property of the n-weight Gauss quadrature rule and the approximation of (5.1) obtained by the n-th iteration of a Lanczos algorithm.

5.1 Orthogonal Polynomials and Krylov Subspaces

Given a matrix $A \in \mathbb{C}^{k\times k}$ and a vector $\mathbf{v} \in \mathbb{C}^k$, we define the n-th *Krylov subspace* generated by A and $\mathbf{v}$ as

$$\mathcal{K}_n(A, \mathbf{v}) = \text{span}\{\mathbf{v}, A\mathbf{v}, \ldots, A^{n-1}\mathbf{v}\}.$$

Let ℓ be the dimension of $\mathcal{K}_n(A, \mathbf{v})$, clearly $\ell \leq n$ and $\ell \leq k$. Moreover, we have the following well-known equivalence.

Lemma 5.1. *Let ℓ be the dimension of $\mathcal{K}_n(A, \mathbf{v})$, with $A \in \mathbb{C}^{k\times k}$ and $\mathbf{v} \in \mathbb{C}^k$ a non zero vector. The following statements are equivalent:*

- *ℓ is the maximal integer such that the dimension of $\mathcal{K}_\ell(A, \mathbf{v})$ is ℓ;*
- *ℓ is the degree of the minimal polynomial of $\mathbf{v}$ with respect to A, i.e. the polynomial p of minimal degree such that $p(A)\mathbf{v} = 0$;*
- *ℓ is the smallest integer for which $\mathcal{K}_\ell(A, \mathbf{v})$ is an A-invariant subspace, i.e. $A\mathbf{w} \in \mathcal{K}_\ell(A, \mathbf{v})$ for every $\mathbf{w} \in \mathcal{K}_\ell(A, \mathbf{v})$.*

Proof. If ℓ is the maximal integer such that the dimension $\mathcal{K}_\ell(A, \mathbf{v})$ is ℓ, then given $\mathbf{w} \in \mathcal{K}_\ell(A, \mathbf{v})$, $A\mathbf{w}$ can be written using a basis of $\mathcal{K}_\ell(A, \mathbf{v})$. Moreover, $\mathcal{K}_n(A, \mathbf{v})$ has dimension n, for every $n < \ell$. Therefore $A^n\mathbf{v} \notin \mathcal{K}_n(A, \mathbf{v})$. Hence the first statement implies the third one.

If ℓ is the smallest integer for which $\mathcal{K}_\ell(A, \mathbf{v})$ is an A-invariant subspace, then

$$A^\ell\mathbf{v} = \sum_{i=0}^{\ell-1} \gamma_i A^i\mathbf{v}.$$

Hence the polynomial $q(x) = x^\ell - \sum_{i=0}^{\ell} \gamma_i x^i$ satisfies $q(A)\mathbf{v} = 0$. Let $\hat{q}$ be a polynomial of degree $n < \ell$ such that $\hat{q}(A)\mathbf{v} = 0$. We can rewrite this equation as

$$A^n\mathbf{v} = \sum_{i=0}^{n-1} \hat{\gamma}_i A^i\mathbf{v}.$$

Thus, every $\mathbf{w} \in \mathcal{K}_{n+1}(A, \mathbf{v})$ can be expressed using a basis of $\mathcal{K}_n(A, \mathbf{v})$. But this contradict the assumption. Hence q is the minimal polynomial of $\mathbf{v}$ with respect to A.

Let ℓ be the degree of the minimal polynomial of $\mathbf{v}$ with respect to A. Since $p(A)\mathbf{v} \neq 0$ for every polynomial p of degree strictly lower than ℓ, vectors

$\mathbf{v}, \ldots, A^n\mathbf{v}$ are linearly independent. Moreover, $\mathcal{K}_{\ell+1}(A, \mathbf{v})$ has dimension $\ell - 1$. Indeed, equation $q(A)\mathbf{v} = 0$ can be expressed as

$$A^\ell \mathbf{v} = \sum_{i=0}^{\ell-1} \gamma_i A^i \mathbf{v}.$$

This ends the proof. □

As remarked in [29, Section 1.1] there is a relation between $\mathcal{K}_n(A, \mathbf{v})$ and $\mathcal{P}_{n-1}$, the subspace of polynomials of degree at most $n - 1$. Indeed, every $\mathbf{u} \in \mathcal{K}_n(A, \mathbf{v})$ can be written as $\mathbf{u} = \alpha_{n-1}A^{n-1}\mathbf{v} + \cdots + \alpha_0$, for some coefficients $\alpha_0, \ldots, \alpha_{n-1} \in \mathbb{C}$. Thus, $\mathbf{u}$ can be associated with the polynomial $p^{(\mathbf{u})}(x) = \alpha_{n-1}x^{n-1} + \cdots + \alpha_0$. Moreover, any basis of $\mathcal{P}_{n-1}$ produces a basis for $\mathcal{K}_n(A, \mathbf{v})$ and

$$\mathcal{K}_n(A, \mathbf{v}) = \{p(A)\mathbf{v} : p \in \mathcal{P}_{n-1}\}.$$

Assuming that ℓ, the dimension of $\mathcal{K}_n(A, \mathbf{v})$, is equal to n, the map $\mathbf{u} \to p^{(\mathbf{u})}$ is an isomorphism between $\mathcal{K}_n(A, \mathbf{v})$ and $\mathcal{P}_{n-1}$. Moreover, we can then define an inner product on $\mathcal{K}_n(A, \mathbf{v})$ given an inner product $\langle \cdot, \cdot \rangle$ on $\mathcal{P}_{n-1}$

$$\langle \mathbf{u}, \mathbf{w} \rangle := \langle p^{(\mathbf{u})}, p^{(\mathbf{w})} \rangle.$$

Now, take the matrix A, the vectors $\mathbf{v}, \mathbf{w}$ and the linear functional defined on $\mathcal{P}_{n-1}$ by

$$\mathcal{L}(p) = \mathbf{w}^* p(A)\mathbf{v}, \quad \text{for } p \in \mathcal{P}_{n-1}.$$

We recall that given a polynomial p we have

$$p(A)^* = \bar{p}(A^*),$$

with $\bar{p}$ the polynomial whose coefficients are the conjugates of the coefficients of p; see (2.2), Chapter 2. Then, for $p, q \in \mathcal{P}_{n-1}$

$$\mathcal{L}(pq) = \mathbf{w}^* p(A)q(A)\mathbf{v} = \hat{\mathbf{w}}^* \hat{\mathbf{v}},$$

with $\hat{\mathbf{v}} = q(A)\mathbf{v} \in \mathcal{K}_n(A, \mathbf{v})$ and $\hat{\mathbf{w}} = \bar{p}(A^*)\mathbf{w} \in \mathcal{K}_n(A^*, \mathbf{w})$,

Then, the orthogonal polynomials $p_0, \ldots, p_{n-1}$ with respect to $\mathcal{L}$ exist if and only if there exist bases $\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}$ and $\mathbf{w}_0 \ldots, \mathbf{w}_{n-1}$ for $\mathcal{K}_n(A, \mathbf{v})$ and $\mathcal{K}_n(A^*, \mathbf{w})$ respectively with the biorthogonality condition

$$\mathbf{w}_i^* \mathbf{v}_j = 0 \quad \text{for } i \neq j, \quad \text{and} \quad \mathbf{w}_i^* \mathbf{v}_i \neq 0, \tag{5.2}$$

for $i, j = 0, \ldots, n$. Indeed, $\mathbf{v}_i = p_i(A)\mathbf{v}$ and $\mathbf{w}_i = \bar{p}_i(A^*)\mathbf{w}$ for $i = 0, \ldots, n-1$.

When A is Hermitian and $\mathbf{v} = \mathbf{w} \neq 0$ we have some important properties related to properties of positive definite linear functionals. We first notice

that $\mathcal{K}_n(A, \mathbf{v}) = \mathcal{K}_n(A^*, \mathbf{w})$. Moreover, $\mathcal{L}(pq)$ is an inner product. Hence, the moments of $\mathcal{L}$ are real and positive and $\mathcal{L}$ is a positive-definite linear functional; see Chapter 1, Definition 1.10. Furthermore, there exists a positive non-decreasing distribution function μ supported on the real axis and having finitely many points of increase such that

$$\mathcal{L}(p) = \int_{\mathbb{R}} p(x) \mathrm{d}\mu(x), \quad \text{for all } p \in \mathcal{P}.$$

Indeed, since A is diagonalizable it can be rewritten as

$$A = Q^* \Lambda Q,$$

where Λ is the diagonal matrix containing the eigenvalues $\lambda_1, \ldots, \lambda_k$ of A and Q is the unitary matrix whose columns are the corresponding eigenvectors. Hence,

$$\mathbf{v}^* p(A)\, \mathbf{v} = \mathbf{v}^* Q^* \Lambda Q\, \mathbf{v} = \mathbf{q}^* p(A) \mathbf{q} = \sum_{i=1}^{k} q_i^2 p(\lambda_i),$$

with $q_1, \ldots, q_k$ the elements of $\mathbf{q} = Q\, \mathbf{v}$. Hence, μ can be defined as

$$\mu_n(x) = \begin{cases} 0, & \text{if } x < \lambda_1 \\ \sum_{i=1}^{j} q_i^2, & \text{if } \lambda_j \leq x < \lambda_{j+1},\ j = 1, \ldots, n-1 \\ \sum_{i=1}^{n} q_i^2 = m_0, & \text{if } \lambda_n \leq x. \end{cases},$$

where $m_0 = \mathcal{L}(x^0) = \mathbf{v}^* \mathbf{v} = ||\mathbf{v}||^2$; see for example [42, Section 7.1]. We refer to Appendix A for the general case of a positive definite linear functional. The orthogonal polynomials $p_0, \ldots, p_{n-1}$ with respect to $\mathcal{L}$ exist and are associated with the orthogonal basis $\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}$ for $\mathcal{K}_n(A, \mathbf{v})$ given by $\mathbf{v}_i = p_i(A)\mathbf{v}$ for $i = 0, \ldots, n-1$.

5.2 Hermitian and non-Hermitian Lanczos Algorithm

In order to find a sequence of polynomials orthonormal with respect to a general linear functional $\mathcal{L}$ we can use the three-term recurrence relation (see (1.12), Chapter 1), that corresponds to the so called *Stieltjes Procedure* (Algorithm 5.2); refer to [29, Algorithm 2.1.3], [34, p. 119], [42, Chapter 7] and [63, Section 3.5]. We remark that in this chapter we always consider algorithms in exact arithmetic.

Algorithm 5.2 (Stieltjes Procedure)**.**
Input: a linear functional $\mathcal{L}$ quasi-definite on $\mathcal{P}_n$
Output: the polynomial $p_0, \ldots, p_n$ orthonormal with respect to $\mathcal{L}$.

Initialize: $p_{-1} = 0, \beta_0 = \sqrt{m_0} = \sqrt{\mathcal{L}(x^0)}, p_0 = 1/\beta_0$.
For $j = 1, 2, \ldots, n$
$\alpha_{j-1} = \mathcal{L}(xp_{j-1}^2(x))$,
$\hat{p}_j(x) = (x - \alpha_{j-1})p_{j-1}(x) - \beta_{j-1}p_{j-2}(x)$,
$\beta_j = \sqrt{\mathcal{L}(\hat{p}_j^2)}$,
$p_j(x) = \hat{p}_j(x)/\beta_j$,
end.

Let A be a $k \times k$ Hermitian matrix and $\mathbf{v}$ be a nonzero vector of dimension k. We are interested in finding an orthogonal basis for $\mathcal{K}_n(A, \mathbf{v})$. Using Algorithm 5.2 we can obtain $p_0, \ldots, p_{n-1}$ orthonormal polynomials with respect to the linear functional $\mathcal{L}$ determined by

$$\mathcal{L}(f) = \mathbf{v}^* f(A)\mathbf{v}, \quad \text{for } f \in \mathcal{P}.$$

Notice that, since A is Hermitian and $\mathbf{v} \neq 0$, $\mathcal{L}$ is positive-definite. Indeed, $\mathcal{L}(p) > 0$ for every nonzero and nonnegative real polynomial from $\mathcal{P}_{2k}$. As we showed in the previous section an orthonormal basis for $\mathcal{K}_n(A, \mathbf{v})$ is given by the vectors $\mathbf{v}_i = p_i(A)\mathbf{v}$ for $i = 0, \ldots, n-1$. Modifying Algorithm 5.2 in order to compute vectors $\mathbf{v}_i$ we obtain the Hermitian Lanczos Algorithm 5.3. This method was introduced by Lanczos in [58, 59] (we also refer to [29, Section 4], [42, Section 4.1], [63, Section 2.4.1] and [66]).
We notice that the algorithm stops before the n-th iteration when $\beta_j = 0$. However, $\beta_j = 0$ if and only if $\mathcal{L}(\hat{p}_j^2) = 0$. Hence, if and only if $\mathcal{L}$ is not quasi-definite on $\mathcal{P}_j$. In addition, $\beta_\ell = 0$ if and only if $\hat{\mathbf{v}}_\ell = 0$. In this case, as shown in the previous section, $\mathbf{v}_0, \ldots, \mathbf{v}_{\ell-1}$ is an orthonormal basis for $\mathcal{K}_\ell(A, \mathbf{v})$. Moreover, $\mathcal{K}_\ell(A, \mathbf{v})$ is an A-invariant subspace since we have

$$A\mathbf{v}_{\ell-1} = \alpha_{\ell-1}\mathbf{v}_{\ell-1} + \beta_{\ell-1}\mathbf{v}_{\ell-2}.$$

Then, by Lemma 5.1, Algorithm 5.3 stops at the ℓ step if and only if $\beta_\ell = 0$ or, equivalently, ℓ is:

- the maximal integer such that the dimension $\mathcal{K}_\ell(A, \mathbf{v})$ is ℓ;

Algorithm 5.3 (Hermitian Lanczos Algorithm)**.**
Input: a Hermitian matrix $A \in \mathbb{C}^{k\times k}$*, a nonzero vector* $\mathbf{v} \in \mathbb{C}^k$*.*
Output: the vectors $\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}$ *orthonormal basis of* $\mathcal{K}_n(A, \mathbf{v})$*.*

Initialize: $\mathbf{v}_{-1} = 0$, $\beta_0 = ||\mathbf{v}|| = \sqrt{\mathbf{v}^*\mathbf{v}}$, $\mathbf{v}_0 = \mathbf{v}/\beta_0$.
For $j = 1, 2, \ldots, n$
$\mathbf{u}_{j-1} = A\mathbf{v}_{j-1} - \beta_{j-1}\mathbf{v}_{j-2}$,
$\alpha_{j-1} = \mathbf{u}_{j-1}^*\mathbf{v}_{j-1}$,
$\hat{\mathbf{v}}_j = \mathbf{u}_{j-1} - \alpha_{j-1}\mathbf{v}_{j-1}$,
$\beta_j = ||\hat{\mathbf{v}}_j||$,
if $\beta_j = 0$ *then stop*,
$\mathbf{v}_j = \hat{\mathbf{v}}_j/\beta_j$,
end.

- the degree of the minimal polynomial of $\mathbf{v}$ with respect to A;
- the smallest integer for which $\mathcal{K}_\ell(A, \mathbf{v})$ is an A-invariant subspace.

The vectors $\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}$ from Algorithm 5.3 satisfy the three-term recurrence relation of the orthogonal polynomials $p_0, \ldots, p_{n-1}$, i.e.

$$\beta_j\mathbf{v}_j = (A - \alpha_{j-1})\mathbf{v}_{j-1} - \beta_{j-1}\mathbf{v}_{j-2}, \quad \text{for } j = 1, \ldots, n.$$

Hence, letting $V_n = [\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}]$ we get

$$AV_n = V_nJ_n + \beta_n\mathbf{v}_n\mathbf{e}_n^T. \quad \text{for } n = 1, 2, \ldots, \ell$$

where J_n is the real Jacobi matrix associated with polynomials $p_0, \ldots, p_{n-1}$. Notice that for $n = \ell$ in the equation above we consider $\mathbf{v}_\ell = \hat{\mathbf{v}}_\ell$. The orthonormality property of $p_0, \ldots, p_{n-1}$ gives $V_n^*V_n = I_n$, with I_n the identity matrix of dimension n. Moreover $V_n^*\mathbf{v}_n = 0$; see Chapter 1, Remark 1.7. Hence, V_n is a unitary transformation such that

$$V_n^*AV_n = J_n.$$

For this reason the Hermitian Lanczos algorithm can be seen as a unitary reduction of a Hermitian matrix to a real Jacobi matrix of lower dimension.

Algorithm 5.4 (non-Hermitian Lanczos Algorithm)**.**
Input: a matrix $A \in \mathbb{C}^{k\times k}$*, two vectors* $\mathbf{v}, \mathbf{w} \in \mathbb{C}^k$ *such that* $\mathbf{w}^*\mathbf{v} \neq 0$*.*
Output: the vectors $\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}$ *and* $\mathbf{w}_0, \ldots, \mathbf{w}_{n-1}$ *bases of* $\mathcal{K}_n(A, \mathbf{v})$ *and* $\mathcal{K}_n(A^*, \mathbf{w})$ *respectively.*

Initialize: $\mathbf{v}_{-1} = \mathbf{w}_{-1} = 0$, $\beta_0 = \delta_0 = 0$,
$\mathbf{v}_0 = \mathbf{v}/||\mathbf{v}||$, $\mathbf{w}_0 = \mathbf{w}/(\mathbf{w}^*\mathbf{v}_0)$.
For $j = 1, 2, \ldots, n$
$\alpha_{j-1} = \mathbf{w}_{j-1}^* A\mathbf{v}_{j-1}$,
$\hat{\mathbf{v}}_j = A\mathbf{v}_{j-1} - \alpha_{j-1}\mathbf{v}_{j-1} - \beta_{j-1}\mathbf{v}_{j-2}$,
$\hat{\mathbf{w}}_j = A^*\mathbf{w}_{j-1} - \bar{\alpha}_{j-1}\mathbf{w}_{j-1} - \bar{\beta}_{j-1}\mathbf{w}_{j-2}$,
$\beta_j = \sqrt{\mathbf{w}_{j-1}^* A\hat{\mathbf{v}}_j}$,
if $\beta_j = 0$ *then stop*,
$\mathbf{v}_j = \hat{\mathbf{v}}_j/\beta_j$,
$\mathbf{w}_j = \hat{\mathbf{w}}_j/\beta_j$,
end.

Now, let A be a $k \times k$ complex matrix, and $\mathbf{v}, \mathbf{w}$ complex vectors of dimension k. We define the linear functional

$$\mathcal{L}(f) = \mathbf{w}^* f(A)\mathbf{v}, \quad \text{for } f \in \mathcal{P}. \tag{5.3}$$

Assuming that $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$, we can compute orthogonal polynomials $p_0, \ldots, p_n$ with respect to $\mathcal{L}$ using Algorithm 5.2. In this case, the vectors

$$\mathbf{v}_j = p_j(A)\mathbf{v}, \quad \mathbf{w}_j = \bar{p}_j(A^*)\mathbf{w}, \quad \text{for } j = 0, \ldots, n,$$

satisfy biorthogonality conditions (5.2). Using these conditions and Algorithm 5.2 we get the non-Hermitian Lanczos Algorithm 5.4. This method was introduced by Lanczos in [58] and [59] (for details we refer to [6, Section 2.7.2], [42, Section 4.2], [63, Section 2.4.2] and [77, Chapter 7]).

We remark that it is possible to obtain different versions of the non-Hermitian Lanczos Algorithm using different orthogonal polynomial sequences with respect to the functional (5.3). In our case, Algorithm 5.4 is based on a family of orthonormal polynomials with respect to the functional (5.3). Indeed, using (1.13) in Chapter 1 gives

$$\alpha_{j-1} = \mathcal{L}(xp_{j-1}^2(x)) = \mathbf{w}_0^* p_{j-1}(A) A p_{j-1}(A)\mathbf{v}_0 = \mathbf{w}_{j-1}^* A \mathbf{v}_{j-1},$$

for $j = 1, \ldots, n$. Moreover, since $\beta_j^2 \mathcal{L}(\hat{p}_j^2(x)) = \mathcal{L}(xp_{j-1}\hat{p}_j(x))$ we get

$$\beta_j = \sqrt{\mathcal{L}(xp_{j-1}(x)\hat{p}_j)} = \sqrt{\mathbf{w}_0^* p_{j-1}(A) A \hat{p}_j(A)\mathbf{v}_0} = \sqrt{\mathbf{w}_{j-1}^* A \hat{\mathbf{v}}_j}, \tag{5.4}$$

for $j = 1, \ldots, n$. As we noticed in Section 1.1, we consider the principal value of the square root.

When $\beta_\ell = 0$ for some $\ell < n$ we say that algorithm 5.4 has a *breakdown*. For a detailed discussion about breakdowns we refer to [72, 8, 11, 45, 71, 46]. If $\hat{p}_\ell$ is such that $\mathcal{L}(p_\ell^2) = 0$, then $\beta_\ell = 0$ and $\mathcal{L}$ is not quasi definite on $\mathcal{P}_\ell$. However, breakdowns can occur even if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$ if we choose another version of the non-Hermitian Lanczos algorithm. In [95, pp. 389–391] Wilkinson showed that breakdowns can arise in the case of matrices with very well conditioned eigenvalues and eigenvectors. Thus, the potential for breakdowns is a specific problem of the non-Hermitian Lanczos Algorithm.

Similarly to the Hermitian case, vectors $\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}$ satisfy the same three-term recurrence relation of $p_0, \ldots, p_{n-1}$, i.e.

$$\beta_j \mathbf{v}_j = (A - \alpha_{j-1})\mathbf{v}_{j-1} - \beta_{j-1}\mathbf{v}_{j-2},$$

for $j = 1, \ldots, n$. Since $\mathbf{w}_j = \bar{p}_j(A^*)\,\mathbf{w}_0$, vectors $\mathbf{w}_0, \ldots, \mathbf{w}_{n-1}$ satisfy the three-term recurrence relation of $p_0, \ldots, p_{n-1}$ with conjugate coefficients, i.e.

$$\beta_j \mathbf{w}_j = (A^* - \bar{\alpha}_{j-1})\mathbf{w}_{j-1} - \bar{\beta}_{j-1}\mathbf{w}_{j-2},$$

for $j = 1, \ldots, n$. Let ℓ be the first breakdown index or let $\ell = k$. For $n = 1, 2, \ldots, \ell$ matrices $V_n = [\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}]$ and $W_n = [\mathbf{w}_0, \ldots, \mathbf{w}_{n-1}]$ satisfy

$$AV_n = V_n J_n + \beta_n \mathbf{v}_n \mathbf{e}_n^T,$$
$$A^* W_n = W_n J_n^* + \bar{\beta}_n \mathbf{w}_n \mathbf{e}_n^T,$$

with J_n the Jacobi matrix associated with polynomials $p_0, \ldots, p_{n-1}$. The biorthogonality conditions (5.2) then give

$$W_n^* V_n = I_n$$
$$W_n^* A V_n = J_n.$$

Therefore, the non-Hermitian Lanczos Algorithm can be seen as a reduction of a matrix to a Jacobi matrix of lower dimension. For a further discussion see [63, Section 2.4.1 and 2.4.2].

Finally, when A, $\mathbf{v}, \mathbf{w}$ are all real we have a real linear functional (5.3). However, since A may not be Hermitian and $\mathbf{w} \neq \mathbf{v}$, the Jacobi matrix obtained by Algorithm 5.4 can be complex. Nevertheless, we can define a variant of the non-Hermitian Lanczos algorithm 5.4, using a different sequence of orthogonal polynomials, which uses only real values in the computations and for the tridiagonal matrix T_n associated with the chosen orthogonal polynomials. We will build the algorithm starting from the orthonormal polynomials sequence, and we will modify the sequence when complex coefficients arise. We first notice that $\alpha_0, p_0(x)$ and $p_{-1}(x)$ are real, then we can proceed by induction. Assume that $p_0, \ldots, p_{j-1}$ are polynomials with real coefficients, and $\alpha_{\ell-1}, \beta_\ell$ are real coefficients for $\ell = 1, \ldots, j-1$. Then by Algorithm 5.2 α_{j-1} and $\hat{p}_j$ are real. While,

$$\beta_j = \sqrt{\mathcal{L}(\hat{p}_j^2)},$$

is complex if and only if $\mathcal{L}(\hat{p}_j^2) < 0$. Moreover, if β_j is not real then it is purely imaginary. Hence, dividing it by the imaginary unit we obtain a real number. Equivalently, we can normalize $\hat{p}_j$ in the following ways

$$\bar{\beta}_j = \sqrt{|\mathcal{L}(\hat{p}_j^2)|}, \qquad \bar{p}_j = \frac{\hat{p}_j}{\bar{\beta}_j}.$$

The orthogonal polynomial $\bar{p}_j$ is not orthonormal since

$$\mathcal{L}(\bar{p}_j^2) = \frac{\hat{p}_j^2}{|\mathcal{L}(\hat{p}_j^2)|} = -1.$$

Algorithm 5.5 (Real non-Hermitian Lanczos Algorithm)**.**
Input: a matrix $A \in \mathbb{R}^{k\times k}$*, two vectors* $\mathbf{v}, \mathbf{w} \in \mathbb{R}^k$ *such that* $\mathbf{w}^*\mathbf{v} \neq 0$*.*
Output: the vectors $\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}$ *and* $\mathbf{w}_0, \ldots, \mathbf{w}_{n-1}$ *bases of* $\mathcal{K}_n(A, \mathbf{v})$ *and* $\mathcal{K}_n(A^*, \mathbf{w})$ *respectively.*

Initialize: $\mathbf{v}_{-1} = \mathbf{w}_{-1} = 0$, $\beta_0 = \delta_0 = 0, \hat{s} = 1, s = 1$,
$\mathbf{v}_0 = \mathbf{v}/||\mathbf{v}||$, $\mathbf{w}_0 = \mathbf{w}/(\mathbf{w}^*\mathbf{v}_0)$.
For $j = 1, 2, \ldots, n$
$\alpha_{j-1} = s \cdot \mathbf{w}_{j-1}^* A\mathbf{v}_{j-1}$,
$\hat{\mathbf{v}}_j = A\mathbf{v}_{j-1} - \alpha_{j-1}\mathbf{v}_{j-1} - \gamma_{j-1}\mathbf{v}_{j-2}$,
$\hat{\mathbf{w}}_j = A^*\mathbf{w}_{j-1} - \alpha_{j-1}\mathbf{w}_{j-1} - \gamma_{j-1}\mathbf{w}_{j-2}$,
$s = sign\,(\mathbf{w}_{j-1}^* A\hat{\mathbf{v}}_j)$,
if $s = 0$ *then stop*,
$\beta_j = \sqrt{|\mathbf{w}_{j-1}^* A\hat{\mathbf{v}}_j|}$,
$\gamma_j = s \cdot \hat{s} \cdot \beta_j$,
$\hat{s} = s$,
$\mathbf{v}_j = \hat{\mathbf{v}}_j/\beta_j$,
$\mathbf{w}_j = \hat{\mathbf{w}}_j/\beta_j$,
end.

Now, let us modify Algorithm 5.2 defining $\beta_j = \sqrt{|\mathcal{L}(\hat{p}_j^2)|}$. Then, (1.13) gives

$$\gamma_j = \frac{\mathcal{L}(xp_{j-1}p_j)}{\mathcal{L}(p_{j-1}^2)} = \frac{\mathcal{L}(p_j^2)}{\mathcal{L}(p_{j-1}^2)}\beta_j = \begin{cases} \beta_j, & \text{if} \quad \mathcal{L}(p_{j-1}^2)\cdot\mathcal{L}(p_j^2) = 1 \\ -\beta_j, & \text{if} \quad \mathcal{L}(p_{j-1}^2)\cdot\mathcal{L}(p_j^2) = -1, \end{cases}$$

$$\alpha_j = \frac{\mathcal{L}(xp_j^2)}{\mathcal{L}(p_j^2)} = \begin{cases} \mathcal{L}(xp_j^2), & \text{if} \quad \mathcal{L}(p_j^2) = 1 \\ -\mathcal{L}(xp_j^2), & \text{if} \quad \mathcal{L}(p_j^2) = -1. \end{cases}$$

Notice that α_j, γ_j are real. Hence, this shows that Algorithm 5.5 involves only real computations and real outputs. Moreover, the tridiagonal matrix $T_n = W_n^* A V_n$ obtained by the first n iterations of the algorithm has sub- and super diagonal elements such that $\beta_j = \pm\gamma_j$ for $j = 1, \ldots, n-1$.

5.3 Lanczos methods and Moment Matching Property

Now we are ready to obtain an approximation of

$$\mathbf{w}^* f(A)\,\mathbf{v}, \tag{5.5}$$

with A a $k \times k$ complex matrix, $\mathbf{v}, \mathbf{w}$ vectors in $\mathbb{C}^k$ such that $\mathbf{w}^*\mathbf{v} \neq 0$, and f a matrix function defined on the spectrum of A (see Chapter 2). Every matrix function can be seen as a matrix polynomial, indeed there exists $p \in \mathcal{P}_k$ depending on A such that $f(A) = p(A)$; see Corollary 2.4. The approximation of (5.5) can be seen as the problem of approximating the linear functional $\mathcal{L} : \mathcal{P}_k \to \mathbb{C}$ defined by

$$\mathcal{L}(x^i) = \mathbf{w}^*(A)^i\,\mathbf{v}, \qquad \text{for } i = 0, \ldots, k. \tag{5.6}$$

In Chapter 4 we introduced the n-weight Gauss quadrature rule (4.6)

$$\mathcal{L}(f) \approx \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j}\, f^{(j)}(\lambda_i),$$

with $\omega_{i,j}$ the weights, λ_i the nodes, and $n = s_1 + \cdots + s_\ell$. If $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$ by Theorem 3.15 and Theorem 4.2 for every $f \in \mathcal{P}_{2n-1}$ we have

$$\mathcal{L}(f) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j}\, f^{(j)}(\lambda_i) = m_0\, \mathbf{e}_1^T f(J_n)\, \mathbf{e}_1,$$

with $m_0 = \mathcal{L}(x^0) = \mathbf{w}^*\mathbf{v}$ and J_n the Jacobi matrix associated with the sequence of polynomial $p_0, \ldots, p_n$ orthonormal with respect to $\mathcal{L}$.

In the previous section we saw that we can compute J_n by the non-Hermitian Lanczos Algorithm 5.4. Assuming that there are no breakdowns in the non-Hermitian Lanczos Algorithm in the first n steps we can compute J_n and hence the approximation of (5.5). Moreover, if a breakdown arises at the ℓ-th iteration, then $\mathcal{L}(p_\ell^2) = 0$. Hence, $\mathcal{L}$ is not quasi-definite on $\mathcal{P}_\ell$. This means that the ℓ-th weight Gauss quadrature rule may have some problems, see Theorems 4.3 and 4.4. This implies that the $(\ell+1)$-st orthogonal polynomial p_ℓ does not exist. In this case the breakdown is known as *true breakdown.*

If A is Hermitian and $\mathbf{v} = \mathbf{w}$ we obtain the same results. However, in this case we can use the n-node Gauss quadrature rule (4.2) for positive definite linear functional $\mathcal{L}$. Hence,

$$\mathcal{L}(f) = \sum_{i=1}^{n} \omega_i\, f(\lambda_i) = m_0\, \mathbf{e}_1^T f(J_n)\, \mathbf{e}_1,$$

where in this case J_n is a real Jacobi matrix obtained using the first n-steps of the Hermitian Lanczos Algorithm. Of course, these are classical well-known results (see [87, Chapters III and XV], [15, Chapter I, Section 6], [35], [42, Section 7.1], [36, Chapter 3.2], [63, Section 3.2], [39] and [40]). In addition, for the Hermitian case it is possible to give upper and lower bounds for the value of (5.5) using Gauss–Radau and Gauss–Lobatto quadrature rules; for more details see [42, Chapter 7]

We conclude this section showing the link between the results presented in this thesis and the Krylov subspaces, through the Vorobyev Method of Moments. Let V_n, W_n and J_n be the outputs of the n-th iteration of Algorithm 5.4 with inputs $A, \mathbf{v}, \mathbf{w}$. By the biorthogonality $W_n^* V_n = I_n$ the oblique projection onto $\mathcal{K}_n(A, \mathbf{v})$ orthogonal to $\mathcal{K}_n(A^*, \mathbf{w})$ is expressed by

$$P_n = V_n W_n^*.$$

Hence, we can define the matrix

$$A_n = P_n A P_n = V_n W_n^* A V_n W_n^* = V_n J_n W_n^*,$$

that is the projection of A onto $\mathcal{K}_n(A, \mathbf{v})$ orthogonally to $\mathcal{K}_n(A^*, \mathbf{w})$. Therefore, we get

$$(A_n)^i = V_n (J_n)^i W_n^*, \quad \text{and} \quad \mathbf{w}^* (A_n)^i \, \mathbf{v} = m_0 \, \mathbf{e}_1^T (J_n)^i \, \mathbf{e}_1, \quad \text{for } i = 0, 1, \ldots.$$

Then, Moment Matching Property gives

$$\mathbf{w}^* A^i \, \mathbf{v} = m_0 \, \mathbf{e}_1^T (J_n)^i \, \mathbf{e}_1 = \mathbf{w}^* (A_n)^i \, \mathbf{v}, \quad \text{for } i = 0, \ldots, 2n-1.$$

Moreover, since P is a projection onto $\mathcal{K}_n(A, \mathbf{v})$ we get

$$\begin{aligned} A_n \mathbf{v} &= A \mathbf{v} \\ A_n (A \mathbf{v}) &= A^2 \mathbf{v} \\ &\vdots \\ A_n (A^{n-2} \mathbf{v}) &= A^{n-1} \mathbf{v} \\ A_n (A^{n-1} \mathbf{v}) &= V_n W_n^* A^n \mathbf{v}. \end{aligned}$$

Or equivalently

$$\begin{aligned} A_n \mathbf{v} &= A \mathbf{v} \\ A_n^2 \mathbf{v} &= A^2 \mathbf{v} \\ &\vdots \\ A_n^{n-1} \mathbf{v} &= A^{n-1} \mathbf{v} \\ A_n^n \mathbf{v} &= V_n W_n^* A^n \mathbf{v}. \end{aligned}$$

These equations are the operator (or vector) moment problem given by Vorobyev; see [91, Chapter VI] and, for the Hermitian case, [91, Chapter III, Sections 2-4, in particular equation (11) p. 54]. In [85] Strakoš shows that using V_n and W_n obtained by the first n steps of the non-Hermitian Lanczos Algorithm we can prove the Moment Matching Property 3.15

$$\mathcal{L}(x^i) = m_0\, \mathbf{e}_1^T (J_n)^i\, \mathbf{e}_1, \qquad \text{for } i = 0, \ldots, 2n-1$$

using the Vorobyev moment problem for a linear functional such that $\mathcal{L}(x^i) = \mathbf{w}^* A^i\, \mathbf{v}$ for $i = 0, \ldots, 2n-1$.

CHAPTER 6

Applications

6.1 Subgraph Centrality for Complex Networks

The analysis of networks has became important in many fields during the last years. In fact, we can use networks to represent many different kinds of relationships between different objects. From the relations between people (social networks), to the interactions between different species in ecology, from the hyperlinks between web sites, to the study of transport routes. For a deeper discussion and more details we refer to [4, 5, 13, 22, 23, 67, 68, 69].

Intuitively a network, or a graph, is a set of objects, called nodes, and links between them, called edges. Usually, we represent a graph with a set of points called the nodes, and when two nodes are connected, there is an arrow (called the edge) from the first to the second node. One of the issues in network analysis is to understand which nodes are the most *important* ones in a network. For this reason we are interested in computing indexes of importance for every node, these are known as *centrality indexes*. Usually, the nodes are then sorted accordingly to their centrality index. Naturally, the meaning of this rank depends on what we consider *important* and, hence, how we define the centrality of a node. For more details we refer to [5]. There are many definitions and many algorithms for computing ranking and centrality indexes (see, e.g., [30, 60, 57, 61, 62, 90]). In this Chapter we consider a particular kind of node centrality, known as *subgraph centrality*, that was introduced by Estrada and Rodríguez-Velázquez in [26]. We refer also to

[26, 3, 25].

We start by recalling some definitions and main properties of graph theory.

Definition 6.1 (Graph)**.** *A* graph G *is an ordered pair of sets* $(V(G), E(G))$ *such that* $V(G)$ *is the* nodes (or vertices) set *and* $E(G) \subset V(G) \times V(G)$ *is the* edges set*.*

The elements of $E(G)$ could be ordered or unordered. We will call a *directed graph* or *digraph* the graph of the first case and an *undirected graph* the second one. An edge (u, v) of a directed graph is usually represented by an arrow from the first node u, the *tail*, to the second one v, the *head*. Two nodes connected by an edge of the graph are said to be *adjacent* nodes? The set of the vertices adjacent to a vertex is called the *neighborhood* of the vertex. Moreover, we say that an edge is *incident* to a node if the node is the tail or the head of the edge.

Definition 6.2 (Degree of a vertex)**.** *The* degree *of a vertex* v *in a graph,* $deg(v)$*, is the number of edges incident to it. In a directed graph we call the* outdegree *of a vertex* v*,* $outdeg(v)$*, the number of edges for which* v *is the tail, while the* indegree *of a vertex* v*,* $indeg(v)$*, is the number of edges for which* v *is the head.*

Definition 6.3 (adjacency matrix)**.** *The* adjacency matrix *of a graph* G *is the matrix* A *such that* $A_{i,j} = 1$ *if* i *and* j *are nodes such that* $(i, j) \in E(G)$ *and* $A_{i,j} = 0$ *elsewhere.*

We remark that an adjacency matrix is symmetric if and only if the graph is undirected. Moreover, in the directed case, the summation of the elements in the i-th row of an adjacency matrix is the outdegree of the i-th node. While the summation of the elements in the j-th column is the indegree of the j-th vertex. A *path* is a sequence of edges $(i_1, i_2), (i_2, i_3), \ldots, (i_{n-1}, i_n)$ of a graph. The following proposition is fundamental for the definition of *subgraph centrality indexes.*

Proposition 6.4. *Let* A *be the incidence matrix of a graph, then*

$$(A^k)_{i,j} = \textit{number of paths of length } k \textit{ from } i \textit{ to } j.$$

Proof. We prove the result by induction. First, we notice that $A_{i,j}$ counts the number of path of length 1 between i and j. Let us assume that $A^{k-1}_{i,\ell}$ is the number of paths of length $k-1$ from the node i to the node ℓ, for $\ell = 1, \ldots, k$. Multiplying the i-th row of A^{k-1} by the j-th column of A means

to add the number of paths of length $k-1$ from the node i to the node ℓ for every node ℓ that is adjacent to the node j. Hence, $A^k_{i,j}$ is the number of paths of length k from i to j. □

A *complex network* is a graph that for our purpose can be thought as a graph obtained from the representation of relationships existent in nature. For example the relations of people in a social network or the mutual citations of the scientific articles in a database. In particular we are interested in two important properties of the adjacency matrix A of a complex network:

- A is a large matrix;
- A is a sparse matrix.

A *sparse* matrix is a matrix such that the number of non-zero elements of the matrix is of the order of $\log(k)$, where k is the order of the matrix.

The degree of a vertex can be seen as a centrality index. Indeed, a vertex adjacent to a many other vertices can be considered more important than a vertex with a low degree. However, it is a local index, since it considers only the neighborhood of the vertex and ignores what is the structure in the other part of the network, or within its neighborhood. Then, to understand the importance of a node in a network it could be interesting to investigate all the possible paths in which that node is involved. For this reason, we will present the following results; see [26, Subsection 7.2.3].

By Proposition 6.4 we can define a subgraph centrality index SC that consists in a weighted sum of the number of closed paths (cycles) passing through a node i, i.e.,

$$\mathrm{SC}(i) = \left(\sum_{\ell=0}^{\infty} \alpha_\ell A^\ell \right)_{i,i}$$

Proposition 2.9 in Chapter 2 shows that, when it converges, the series

$$f(A) = \sum_{\ell=0}^{\infty} \alpha_\ell A^\ell \tag{6.1}$$

is a matrix function, see Definition 2.2. Moreover, f is the function defined by the scalar series

$$f(x) = \sum_{\ell=0}^{\infty} \alpha_\ell x^\ell,$$

for every x for which it converges. For more details see Chapter 2. Hence, we have to choose the weights α_ℓ so that the series is convergent for the adjacency matrices we are considering. Moreover, the choice must have an

interpretation with respect to the main idea of counting the number of paths passing through the node i. For these reasons we need a decreasing sequence of weights $\alpha_0, \alpha_1, \ldots$. In particular, we can consider the following two choices:

- $\alpha_\ell = 1/\ell!$, which gives $f(A) = \exp(A) = e^A$, where e^A is the exponential of a matrix (see Chapter 2);
- $\alpha_\ell = \alpha^\ell$ with $0 < \alpha < 1/|\lambda_k|$, which gives $f(A) = (I - \alpha A)^{-1}$, the resolvent of a matrix, with λ_k the eigenvalue of A of maximal modulus.

Thus, we can compute the centrality index using a bilinear form

$$SC(i) = \mathbf{e}_i^T f(A)\mathbf{e}_i, \tag{6.2}$$

where $\mathbf{e}_i$ is the i-th vector of the canonical basis.

Then to approximate (6.2) we will use Lanczos algorithms, that we illustrated in Chapter 5. Indeed, from the j-th iteration of the real non-Hermitian Lanczos algorithm 5.5 we obtain a tridiagonal matrix T_j, of dimension j, such that

$$SC(i) = \mathbf{e}_i^T f(A)\mathbf{e}_i \approx \mathbf{e}_1^T f(T_j)\mathbf{e}_1,$$

see Section 5.2. Hence, we reduce the problem from the approximation of the matrix function of a $k \times k$ matrix to the approximation of the matrix function of a $j \times j$ matrix. This was first proposed by Benzi in [3] for undirected networks, i.e., for symmetric (Hermitian) adjacency matrices. Hence, our purpose is to use the non-Hermitian Lanczos algorithm for directed graphs, since we are dealing with non-symmetric adjacency matrices.

6.2 Numerical Experiments

In this section we present some preliminary results. We are still studying the problem and many questions arising from the first experiments still need an answer. We want to approximate some diagonal entries of the matrix $\exp(A)$, with A the adjacency matrix of a directed network. We will consider the matrix A from a data set of small web graphs consisting of web sites on various topics. We use one of the matrices used in the experiments in [3] obtained following the procedure in [56] with query *death penalty*; see in particular [88, Section 6.1]. This is a real non-symmetric 1850×1850 matrix with 7363 nonzero elements. Indeed, it represents a directed graph with 1850 nodes and 7363 edges. Every node corresponds to a web page and every edge is a hyperlink from a page to another one. This kind of networks is known as *hyperlink graph*. We compute the ten greatest entries in the diagonal

of $\exp(A)$ and we compare the results with those of the `expm` function of `Matlab`. All the experiments were performed using `Matlab 7.12.0`. We used Algorithm 5.5 (Real non-Hermitian Lanczos Algorithm) to compute a real tridiagonal matrix T_j, with j the last iteration of the method. Then we obtained the approximation by

$$\mathbf{e}_1^T \exp(T_j)\mathbf{e}_1,$$

in which we use `expm` Matlab function to estimate the exponential of T_j.

Using Algorithm 5.5 for computing the diagonal entries of $\exp(A)$ a problem arise. Let $\mathbf{v}_0 = \mathbf{w}_0 = \mathbf{e}_i$ be the input vector of the algorithm. Then from the first iteration we get $\alpha_0 = A_{i,i}$, the i-th element of the diagonal of the input $k \times k$ real matrix A. Hence

$$\hat{\mathbf{v}}_1 = A\mathbf{e}_i - A_{i,i}\mathbf{e}_i = \mathbf{a}_i - A_{i,i}\mathbf{e}_i,$$

with $\mathbf{a}_i$ the i-th column of A. Therefore, denoting $\hat{\mathbf{a}}_i$ the vector obtained by the i-th row of A, we obtain

$$\mathbf{w}_0^T A\,\hat{\mathbf{v}}_1 = \mathbf{e}_i^T A(\mathbf{a}_i - A_{i,i}\mathbf{e}_i) = \hat{\mathbf{a}}_i^T\mathbf{a}_i - (A_{i,i})^2,$$

which is different from zero if and only if there exist an index $j \neq i$ such that $A_{i,j} = A_{j,i} = 1$. Since A is a sparse matrix this is very unlikely. Thus we often have a breakdown at the first step of the algorithm.

To overcome this problem we need to use some non-sparse vector $\mathbf{v}$ as input. We propose the following procedure. Let us define the vectors:

$$\mathbf{e} = (1, \ldots, 1)^T, \quad \mathbf{v} = \frac{\mathbf{e}_i + \mathbf{e}}{\sqrt{k+3}}, \quad \mathbf{w} = \frac{\sqrt{k+3}\,\mathbf{e}_i}{2}.$$

Then, $\mathbf{w}^T\mathbf{v} = 1$ and

$$2\,\mathbf{w}^T \exp(A)\,\mathbf{v} = \mathbf{e}_i^T \exp(A)\,\mathbf{e}_i + \mathbf{e_i}^T \exp(A)\,\mathbf{e},$$

and so we can compute $\mathbf{e}_i^T \exp(A)\,\mathbf{e}_i$ subtracting an approximation of $\mathbf{e_i}^T \exp(A)\,\mathbf{e}$ to an approximation of $2\,\mathbf{w}^T \exp(A)\,\mathbf{v}$.

The approximation in Table 6.1 are obtained using two times Algorithm 5.5, hence, the last column represents the number of iterations necessary to compute respectively $2\,\mathbf{w}^T \exp(A)\,\mathbf{v}$ and $\mathbf{e_i}^T \exp(A)\,\mathbf{e}$. We stop the algorithm at the 10-th iteration for both approximations, with some exceptions we will explain. Moreover, we stop the algorithm at the j-th iteration, with $j < 10$, if $\beta_j < 1e - 10$.

As we can see in Table 6.1, the approximation is good for the first ten values sorted from the biggest to the lowest. However, not all the elements

Table 6.1: First ten entries of the diagonal of $\exp(A)$, with A adjacency matrix of the *death penalty* hyperlink graph

Index	value	err abs	err rel	n. it
1632	2.56307786e+03	3.85171006e-10	1.50276747e-13	10 + 10
1671	7.21827144e+02	2.33171704e-10	3.23029836e-13	10 + 10
1653	5.38668944e+02	2.88764567e-11	5.36070569e-14	10 + 10
1662	4.70964536e+02	9.89075488e-11	2.10010608e-13	10 + 10
552	2.20355022e+02	1.85644921e-09	8.42481006e-12	10 + 10
1651	1.91520312e+02	1.77095671e-10	9.24683494e-13	10 + 10
1640	1.60758638e+02	8.93010110e-11	5.55497432e-13	10 + 10
1639	1.51747810e+02	1.82467374e-11	1.20243826e-13	11 + 11
1638	1.51747810e+02	7.03437308e-11	4.63556809e-13	10 + 10
1641	1.27189646e+02	2.72720512e-10	2.14420372e-12	10 + 10

of the diagonal of $\exp(A)$ were well computed. Indeed, in 8 cases we have a `NaN`, not a number answer, i.e. some divisions by zero occurred.

Finally, in Figure 6.1 we plot the relative error of the approximation of the 1632-nd entry of the exponential matrix stopping the procedure at the i-th iteration, for $i = 1, \dots, 10$. As we can see, the value converges to the one obtained by the `expm` function of Matlab. However, at the fourth iteration we have a wrong result which seems to not influence the following iterations behavior. This is why sometimes we had to add an iteration in order to have a good approximation. It seems that in some isolated iterations one of the approximated values diverges, for then converging again in the following steps. In our opinion, this could be linked with some not well-conditioned tridiagonal matrices obtained by the algorithm. Indeed, if the tridiagonal matrix T_{j-1} is not well conditioned we could have problems with the approximation of $\exp(T_{j-1})$. Nevertheless, since the tridiagonal matrix T_j obtained by the following iteration has different eigenvalues (see Theorem 3.19), then we obtain a better result for the evaluation of $\exp(T_j)$. We hope to give a better interpretation of the phenomenon in future works.

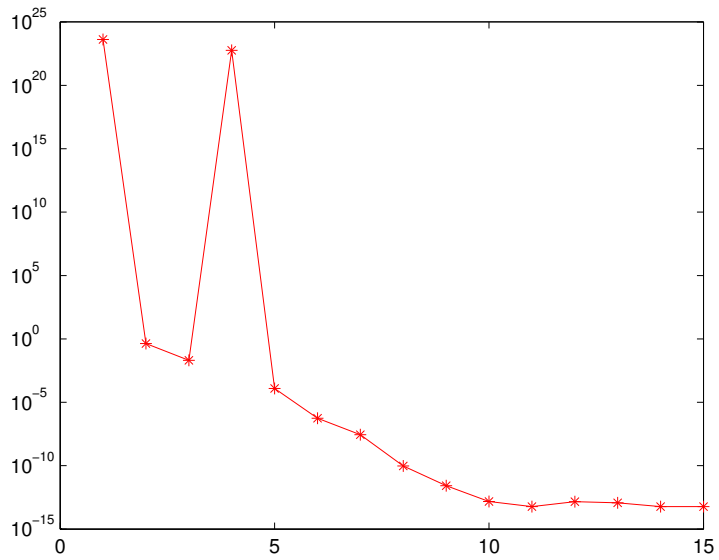


Figure 6.1: Relative errors at every iteration for the computation of the 1632 entry of the diagonal of $\exp(A)$, with A the adjacency matrix of the *death penalty* hyperlink graph. We remark that at every iteration we use two times Algorithm 5.5.

CHAPTER A

The Representation Theorem

Let $\mathcal{L}$ be a linear functional defined on $\mathcal{P}$, the space of polynomials, and let $m_0, m_1, m_2, \ldots$ its moments; see (1.1) Chapter 1. We want to show the equivalence between positive definite linear functionals (see Definition 1.10, Chapter 1) and integrals with respect to some distribution functions.

Theorem A.1. *The linear functional $\mathcal{L}$ is positive definite on $\mathcal{P}_k$ if and only if there exists a positive non-decreasing distribution function μ supported on the real axis such that*

$$\mathcal{L}(p) = \int_{\mathbb{R}} p(x)\, d\mu(x), \qquad \textit{for all } p \in \mathcal{P}_{2k}. \tag{A.1}$$

Proof. Le μ be the measure defined in the theorem. Since μ is positive

$$\int_{\mathbb{R}} p(x) \mathrm{d}\mu(x) > 0,$$

for every nonzero and nonnegative polynomial p. By Theorem 1.12 of Chapter 1 the integral is a positive definite linear functional on $\mathcal{P}$.

Conversely, let $\mathcal{L}$ be positive definite on $\mathcal{P}_k$. For $n = 1, \ldots, k$ the classical n-node Gauss quadrature formula (4.2) gives the relation

$$m_j = \mathcal{L}(x^j) = \sum_{i=1}^{n} \omega_i \lambda_i^j, \qquad \text{for } j = 0, \ldots, 2n-1 \tag{A.2}$$

with $\omega_1, \ldots, \omega_n$ positive weights and $\lambda_1 < \cdots < \lambda_n$ distinct real nodes; see Section 4.1.

Let $n \leq k$ and define the non-decreasing distribution functions μ_n as

$$\mu_n(x) = \begin{cases} 0, & \text{if } x < \lambda_1 \\ \sum_{i=1}^{\ell} \omega_i, & \text{if } \lambda_\ell \leq x < \lambda_{\ell+1},\ \ell = 1, \ldots, n-1 \\ \sum_{i=1}^{n} \omega_i = m_0, & \text{if } \lambda_n \leq x. \end{cases} \tag{A.3}$$

Clearly, μ_n is a bounded, right continuous step function. In addition its points of increase are $\lambda_1, \ldots, \lambda_n$ and the jumps at λ_i are ω_i, for $i = 1, \ldots, n$. Hence

$$\int_{\mathbb{R}} p \, \mathrm{d}\mu_n = \sum_{i=1}^{n} \omega_i p(\lambda_i^j), \qquad \text{for any } p \in \mathcal{P}_{2n-1}. \tag{A.4}$$

For $n = k$ we obtain the desired measure. □

If $\mathcal{L}$ is positive definite on $\mathcal{P}$ and we want that (A.1) holds for every p from $\mathcal{P}$, then we need to complete the previous proof; see Chapter II, sections 1,2 and 3 of [15]. We begin recalling some convergence theorems.

Theorem A.2. *Let $f_0, f_1, \ldots$ be a sequence of real functions defined on a countable set E. If for every x from E $f_0(x), f_1(x), \ldots$ is a bounded sequence, then there exists a subsequence that converges for every $x \in E$.*

Proof. Let $E = \{x_1, x_2, \ldots\}$. The sequence $f_0(x_1), f_1(x_1), f_2(x_1), \ldots$ is a bounded sequence of real numbers, hence there exists a subsequence $f_0^{(1)}, f_1^{(1)}, \ldots$ which is convergent for $x = x_1$. Moreover, $f_0^{(1)}(x_2), f_1^{(1)}(x_2), \ldots$ is a bounded sequence, hence there exist $f_0^{(2)}, f_1^{(2)}, \ldots$ a subsequence of $f_0^{(1)}, f_1^{(1)}, \ldots$ that is convergent in $x = x_2$. Repeating the argument and defining $f_n^{(0)} \equiv f_n$ for $n = 0, 1, \ldots$ gives

1. there exists $f_0^{(k)}, f_1^{(k)}, \ldots$, a subsequence of $f_0^{(k-1)}, f_1^{(k-1)}, \ldots$;
2. $f_0^{(k)}(x), f_1^{(k)}(x), \ldots$ is convergent for $x \in E_k = \{x_1, \ldots, x_k\}$.

With a little care in order to preserve the relative order of the terms, by (1) the diagonal sequence $f_0^{(0)}, f_1^{(1)}, f_2^{(2)}, \ldots$ is a subsequence of $f_0, f_1, \ldots$. Moreover, $f_k^{(k)}, f_{k+1}^{(k+1)}, \ldots$ is a subsequence of $f_0^{(k)}, f_1^{(k)}$. Thus, by property (2) above $f_0^{(0)}, f_1^{(1)}, f_2^{(2)}, \ldots$ converges for $x \in E = \bigcup_{k=0}^{\infty} E_k$. □

If stated in terms of function of bounded variation the following theorem is known as *Helly's Selection Principle* or *Theorem of Choice*. As suggested by Chihara [15, p. 53] for our purpose it is enough to state it for non-decreasing functions.

Theorem A.3. *Let $\mu_0, \mu_1, \mu_2, \dots$ be a uniformly bounded sequence of non-decreasing functions defined on $(-\infty, +\infty)$. There exists a subsequence of $\mu_0, \mu_1, \mu_2, \dots$ which converges on $(-\infty, +\infty)$ to a bounded, non-decreasing function μ.*

Proof. By Theorem A.2 there exists a subsequence $\mu_{t_0}, \mu_{t_1}, \mu_{t_2}, \dots$ that is convergent on $\mathbb{Q}$. If we define on $\mathbb{Q}$ a function $\hat{\mu}$ such that

$$\hat{\mu}(x) = \lim_{i \to \infty} \mu_{t_i}, \qquad \text{for every } x \in \mathbb{Q},$$

then $\hat{\mu}$ is bounded and non-decreasing on $\mathbb{Q}$. Let us extend the domain of $\hat{\mu}$ on $\mathbb{R}$. We define

$$\hat{\mu}(x) = \sup\{\hat{\mu}(y) \,|\, y \in \mathbb{Q}, y < x\}, \qquad \text{for every } x \in \mathbb{R} \setminus \mathbb{Q},$$

so that $\hat{\mu}$ is bounded and non-decreasing on $\mathbb{R}$. Now we show that the subsequence converges to $\hat{\mu}(x)$ at all the points x of continuity of $\hat{\mu}$ from $\mathbb{R}$. Suppose $\hat{\mu}$ continuous at a point $x \in \mathbb{R} \setminus \mathbb{Q}$. By the density of $\mathbb{Q}$ in $\mathbb{R}$ given $\epsilon_1, \epsilon_2 > 0$ there exist $x_1, x_2 \in \mathbb{Q}$ such that $x_1 < x < x_2$ and

$$\hat{\mu}(x) - \epsilon_1 \leq \hat{\mu}(x_1)$$

$$\hat{\mu}(x_2) \leq \hat{\mu}(x) + \epsilon_2.$$

Moreover,

$$\mu_{t_i}(x_1) \leq \mu_{t_i}(x) \leq \mu_{t_i}(x_2).$$

Hence,

$$\hat{\mu}(x_1) \leq \liminf_{i \to \infty} \mu_{t_i}(x) \leq \limsup_{i \to \infty} \mu_{t_i}(x) \leq \hat{\mu}(x_2).$$

Therefore

$$\hat{\mu}(x) - \epsilon_1 \leq \liminf_{i \to \infty} \mu_{t_i}(x) \leq \limsup_{i \to \infty} \mu_{t_i}(x) \leq \hat{\mu}(x) + \epsilon_2,$$

hence $\mu_{t_0}, \mu_{t_1}, \mu_{t_2}, \dots$ converges to $\hat{\mu}$ at all its points of continuity. However, since $\hat{\mu}$ is non-decreasing, the set of its points of discontinuity D is a countable set. Using Theorem A.2 to the subsequence $\mu_{t_0}, \mu_{t_1}, \mu_{t_2}, \dots$ and the set D we obtain a subsequence of $\mu_{t_0}, \mu_{t_1}, \mu_{t_2}, \dots$ that converges on D, and hence on $\mathbb{R}$, to a bounded non-decreasing function μ. Clearly, $\mu \equiv \hat{\mu}$ on $\mathbb{R} \setminus D$. □

We now state *Helly's second theorem*, as done for the previous one we consider only non-decreasing functions.

Theorem A.4. *Let $\mu_0, \mu_1, \mu_2, \ldots$ be a uniformly bounded sequence of non-decreasing functions defined on a compact interval $[a, b]$. If the sequence converges to a limit function μ, then*

$$\lim_{n\to\infty} \int_a^b f \, d\mu_n = \int_a^b f \, d\mu,$$

for every real function f continuous on $[a, b]$.

Proof. Since $\mu_0, \mu_1, \mu_2, \ldots$ is uniformly bounded there exists $M > 0$ such that

$$0 \le \mu(b) - \mu(a) \le M \text{ and } 0 \le \mu_n(b) - \mu_n(a) \le M, \text{ for } n = 0, 1, 2, \ldots.$$

The function f is real and continuous on $[a, b]$, hence it is uniformly continuous on $[a, b]$. Then, given $\epsilon > 0$ there exists a partition $P_\epsilon = \{x_0, x_1, \ldots, x_\nu\}$ of $[a, b]$ for which

$$|f(\tilde{x}) - f(\tilde{\tilde{x}})| < \epsilon, \quad \text{for } \tilde{x}, \tilde{\tilde{x}} \in [x_{i-1}, x_i], \quad \text{for } i = 1, \ldots, \nu.$$

Let us define

$$\Delta_i\mu = \mu(x_i) - \mu(x_{i-1}), \quad \text{and} \quad \Delta_i\mu_n = \mu_n(x_i) - \mu_n(x_{i-1}),$$

for $n = 0, 1, \ldots$ and $i = 0, \ldots, \nu$. We remark that we use Δ_i as the forward difference just in this proof, since in the rest of Part I it is the Hankel determinant. Fixing $y_i \in [x_{i-1}, x_i]$ by the mean value theorem for Stieltjes integrals we get

$$\int_{x_{i-1}}^{x^i} f \, \mathrm{d}\mu - f(y_i)\Delta_i\mu = (f(\tilde{y}_i) - f(y_i))\Delta_i\mu,$$

for some $\tilde{y}_i \in [x_{i-1}, x_i]$.

Summing over i gives

$$\begin{aligned}\left| \int_a^b f \, \mathrm{d}\mu - \sum_{i=0}^{\nu} f(y_i)\Delta_i\mu \right| &\le \sum_{i=0}^{\nu} |(f(\tilde{y}_i) - f(y_i))| \, \Delta_i\mu \\ &< \epsilon \sum_{i=0}^{\nu} \Delta_i\mu \le \epsilon M.\end{aligned}$$

Repeating the same argument we obtain

$$\left| \int_a^b f \, \mathrm{d}\mu_n - \sum_{i=0}^{\nu} f(y_i)\Delta_i\mu_n \right| < \epsilon M, \quad \text{for } n = 0, 1, \ldots.$$

Thus

$$\left|\int_a^b f\,\mathrm{d}\mu - \int_a^b f\,\mathrm{d}\mu_n\right| \le \left|\int_a^b f\,\mathrm{d}\mu - \sum_{i=0}^{\nu} f(y_i)\Delta_i\mu\right| + \left|\sum_{i=0}^{\nu} f(y_i)(\Delta_i\mu - \Delta_i\mu_n)\right| + \left|\int_a^b f\,\mathrm{d}\mu_n - \sum_{i=0}^{\nu} f(y_i)\Delta_i\mu_n\right| < 2\epsilon M + \sum_{i=0}^{\nu} |f(y_i)||(\Delta_i(\mu-\mu_n))|.$$

Fixing the partition P_ϵ we get $\lim_{n\to\infty}\Delta_i(\mu-\mu_n)=0$ for $i=0,\ldots,\nu$. Hence

$$\limsup_{n\to\infty}\left|\int_a^b f\,\mathrm{d}\mu - \int_a^b f\,\mathrm{d}\mu_n\right| < 2\epsilon M$$

ends the proof. □

If $\mathcal{L}$ is positive definite on $\mathcal{P}$, then (A.2) stands for $n=0,1,2,\ldots$. Equation (A.3) defines the sequence $\mu_1,\mu_2,\ldots$. Again, μ_n is a bounded, right continuous, non-decreasing step function and its points of increase are $\lambda_1,\ldots,\lambda_n$. Since $0\le\mu_n(x)\le m_0$ for every $x\in\mathbb{R}$ and $n=1,2,\ldots$, $\mu_1,\mu_2,\ldots$ is a uniformly bounded sequence. Then, Theorem A.3 shows that there exists a subsequence $\mu_{t_1},\mu_{t_2},\ldots$ of the sequence $\mu_1,\mu_2,\ldots$ which converges on $(-\infty,+\infty)$ to a bounded, non-decreasing function μ. Denoting $\phi_i=\mu_{t_i}$ for $i=1,2,\ldots$ and using (A.4) gives

$$\int_{\mathbb{R}} x^j\,\mathrm{d}\phi_i = m_j, \quad \text{for } t_i \ge \frac{j+1}{2}.$$

Theorem A.4 implies

$$\lim_{i\to\infty}\int_a^b x^j\,\mathrm{d}\phi_i = \int_a^b x^j\,\mathrm{d}\phi, \quad \text{for } j=0,1,\ldots, \tag{A.5}$$

for every compact interval $[a,b]$. Setting $a<0<b$ and $n_i>\frac{j+1}{2}$ gives

$$\left|m_j - \int_a^b x^j\,\mathrm{d}\phi\right| = \left|\int_{\mathbb{R}} x^j\,\mathrm{d}\phi_i - \int_a^b x^j\,\mathrm{d}\phi\right| \le \left|\int_{-\infty}^a x^j\,\mathrm{d}\phi_i\right| + \left|\int_b^{+\infty} x^j\,\mathrm{d}\phi_i\right| + \left|\int_a^b x^j\,\mathrm{d}\phi_i - \int_a^b x^j\,\mathrm{d}\phi\right|$$

However, we get

$$\left|\int_b^{+\infty} x^j\,\mathrm{d}\phi_i\right| = \left|\int_b^{+\infty}\frac{x^{2j+2}}{x^{j+2}}\,\mathrm{d}\phi_i\right| \le b^{-(j+2)}\left|\int_b^{+\infty} x^{2j+2}\,\mathrm{d}\phi_i\right| \le b^{-(j+2)}m_{2j+2},$$

since $x^{-(j+2)} \leq b^{-(j+2)}$ for $x \geq b$, and $x^{2j+2} \geq 0$ for $x \in \mathbb{R}$. Similarly we have

$$\left| \int_{-\infty}^{a} x^j \, \mathrm{d}\phi_i \right| \leq |a|^{-(j+2)} m_{2j+2}.$$

Hence, we obtain

$$\left| m_j - \int_a^b x^j \, \mathrm{d}\phi \right| \leq \left| \int_a^b x^j \, \mathrm{d}\phi_i - \int_a^b x^j \, \mathrm{d}\phi \right| + (|a|^{-(j+2)} + b^{-(j+2)}) m_{2j+2}.$$

For $i \to \infty$ by (A.5) we get

$$\left| m_j - \int_a^b x^j \, \mathrm{d}\phi \right| \leq (|a|^{-(j+2)} + b^{-(j+2)}) m_{2j+2}.$$

Letting $a \to -\infty$ and $b \to +\infty$ finishes the proof.

Bibliography

[1] N.I. Akhiezer, The classical moment problem and some related questions in analysis, Oliver & Boyd, Edinburgh, 1965.

[2] B. Beckermann, Complex Jacobi matrices, J. Comput. Appl. Math. 127 (2001) 17–65.

[3] M. Benzi, E. Estrada, C. Klymko, Ranking hubs and authorities using matrix functions, Linear Algebra Appl. 438 (2013) 2447–2474.

[4] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: structure and dynamics, Phys. Rep. 424 (2006) 175–308.

[5] U. Brandes, T. Erlebach, eds., Network analysis: Methodological foundations, Lecture Notes in Computer Science Vol. 3418, Springer, New York, 2005.

[6] C. Brezinski, Padè-type approximation and general orthogonal polynomials, International series of numerical mathematics, Birkhäuser, 1980.

[7] C. Brezinski, M. Redivo Zaglia, Breakdowns in the computation of orthogonal polynomials, in: A. Cuyt (Ed.), Nonlinear numerical methods and rational approximation, Kluwer, Dordrecht, 1994, pp. 49–59.

[8] C. Brezinski, M. Redivo Zaglia, H. Sadok, Avoiding breakdown and near-breakdown in Lanczos type algorithm, Numer. Algorithms 1 (1991) 261–284.

[9] C. Brezinski, M. Redivo Zaglia, H. Sadok, Breakdowns in the implementation of the Lanczos method for solving linear systems, Comput. Math. Appl. 33 (1997) 31–44.

[10] C. Brezinski, M. Redivo Zaglia, H. Sadok, A review of formal orthogonality in Lanczos-based methods, J. Comput. Appl. Math. 140 (2002) 81–98.

[11] C. Brezinski, H. Sadok, Avoiding breakdown in the CGS algorithm, Numer. Algorithms 1 (1991) 199–206.

[12] A. Buchheim, An extension of a theorem of Professor Sylvester relating to matrices, Phil. Mag. (5)22 (1886) 173–174.

[13] G. Caldarelli, Scale-free networks, Oxford University Press, Oxford, UK, 2007.

[14] L. Chakalov, General quadrature formulae of Gaussian type, Bulgar. Akad. Nauk Izv. Mat. Inst. 1 (1954) 67–84.

[15] T.S. Chihara, An introduction to orthogonal polynomials, Gordon and Breach, New York, 1978.

[16] E.B. Christoffel, Über die Gaußische Quadratur und eine Verallgemeinerung derselben, J. Reine Angew. Math. 55 (1858) 61–82. Reprinted in: Gesammelte mathematische Abhandlungen, vol. 1, B.G. Teubner, Leipzig, 1910, pp. 65–87.

[17] E.B. Christoffel, Sur une classe particulière de fonctions entières et de fractions continues, Ann. Mat. Pura Appl. 8 (1877) 1–10. Reprinted in: Gesammelte mathematische Abhandlungen, vol. 2, B.G. Teubner, Leipzig, 1910, pp. 42–50.

[18] T.C. Chu Moody, G.H. Golub, Inverse eigenvalue problems, theory, algorithms, and applications, Oxford University Press, New York, 2005.

[19] M. Cipolla, Sulle matrice espressione analitiche di un'altra, Rend. Circ. Mat. di Palermo 56 (1932) 144–15.

[20] B.D. Craven, Complex symmetric matrices, J. Aust. Math. Soc. 10 (1969) 341–354.

[21] D.Ž. Djoković, Eigenvectors obtained from the adjoint matrix, Aequat. Math. 2 (1969) 94–97.

[22] E. Estrada, The Structure of complex networks: Theory and applications, Oxford University Press, Oxford, 2011.

[23] E. Estrada, M. Fox, D. Higham, and G.-L. Oppo, eds., Network science. Complexity in nature and technology, Springer, New York, 2010.

[24] E. Estrada, N. Hatano, M. Benzi, The physics of communicability in complex networks, Phys. Rep. 514 (2012) 89–119.

[25] E. Estrada, D.J. Higham, Network properties revealed through matrix functions, SIAM Rev. 52(4) (2010) 696–714.

[26] E. Estrada, J.A. Rodríguez-Velázquez, Subgraph centrality in complex networks, Phys. Rev. E 71(5) (2005) 056103

[27] J. Favard, Sur les polynômes de Tchebicheff, C. R. Acad. Sci., Paris, 200 (1935) 2052–2053.

[28] L. Fantappiè, Le calcul des matrices, Comptes Rendus 186 (1928) 619–621.

[29] B. Fischer, Polynomial based iteration methods for symmetric linear systems, Wiley-Teubner Series Advances in Numerical Mathematics, John Wiley & Sons Ltd., Chichester, 1996.

[30] M. Franceschet, PageRank: Standing on the shoulders of giants, Comm. ACM 54 (2011) 92–101.

[31] R.W. Freund, M. Hochbruck, Gauss quadratures associated with the Arnoldi process and the Lanczos algorithm, in: M.S. Moonen, G.H. Golub, B.L.R. De Moor (Eds.), Linear Algebra for Large Scale and Real-Time Application, Kluwer, Dordrecht, The Netherlands, 1993, pp. 377–380.

[32] F.R. Gantmakher, The Theory of matrices 1,2, Chelsea Publishing Co., New York, 1959.

[33] C.F. Gauss, Methodus nova integralium valores per approximationem inveniendi, Commentationes Societatis Regiae Scientiarum Gottingensis (1814) 39–76. Reprinted in: Werke, vol. 3, Göttingen, 1876, pp. 163–196.

[34] W. Gautschi, A survey of Gauss-Christoffel quadrature formulae, in: P. Butzer, F. Fehér (Eds.), E.B. Christoffel: The influence of his work in mathematics and the physical sciences, Birkhäuser, Basel, 1981, pp. 72–147.

[35] W. Gautschi, Orthogonal polynomials: Computation and approximation, Oxford University Press, Oxford, 2004.

[36] W. Gautschi, Numerical analysis, 2nd ed., Birkhäuser, Boston, 2012.

[37] G. Giorgi, Sulle funzioni delle matrice, Atti Accad. Lincei Rend. (6)8 (1928) 3–8.

[38] W. Givens, A method of computing eigenvalues and eigenvectors suggested by classical results on symmetric matrices, in: Simultaneous Linear Equations and the Determination of Eigenvalues, National Bureau of Standards Applied Mathematics Series, 29, U.S. Government Printing Office, Washington, D.C., 1953, pp. 117–122.

[39] G.H. Golub, Bounds for matrix moments, Rocky Mnt. J. Math. 4 (1974) 207–211.

[40] G.H. Golub, Matrix computation and the theory of moments, in: Proceedings of the international congress of mathematicians, Zürich, Switzerland 1994, Birkhäuser, Basel, 1995, pp. 1440–1448.

[41] G.H. Golub, J. Kautsky, Calculation of Gauss quadratures with multiple free and fixed knots, Numer. Math. 41 (1983) 147–163.

[42] G.H. Golub, G.A. Meurant, Matrices, moments, and quadrature with applications, Princeton University Press, Princeton, N.J., 2010.

[43] G.H. Golub, J.H. Welsch, Calculation of Gauss quadrature rules, Math. Comput. 23 (1969) 221–230.

[44] W.B. Gragg, Matrix interpretations and applications of the continued fraction algorithm, Rocky Mountain J. Math. 4 (1974) 213–225.

[45] M.H. Gutknecht, A completed theory of the unsymmetric Lanczos process and related algorithms. I, SIAM J. Matrix Anal. Appl. 13 (1992) 594–639.

[46] M.H. Gutknecht, A completed theory of the unsymmetric Lanczos process and related algorithms. II, SIAM J. Matrix Anal. Appl. 15 (1994) 15–58.

[47] K. Hensel, Über Potenzreihen von Matrizen, J. Reine Angew. Math. 155 (1926) 107–110.

[48] R.A. Horn, C.R. Johnson, Matrix analysis, Cambridge University Press, Cambridge, UK, 1985.

[49] E.D. Hellinger, O. Toeplitz, Zur Einordnung der Kattenbruchtheorie in die Theorie der quadratischen Formen von unendlichvielen Veränderlichen, J. Reine Angew. Math. 144 (1914) 212-238.

[50] E.D. Hellinger, H.S. Wall, Contributions to the analytic theory of continued fractions and infinite matrices, Ann. Math. 44 (1943) 103–127.

[51] N.J. Higham, Functions of matrices: Theory and computation. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.

[52] A.S. Householder, The theory of matrices in numerical analysis, Dover, New York, 1964.

[53] E. Isaacson, H.B. Keller, Analysis of numerical methods, J. Wiley & Sons Inc., New York, 1966.

[54] C.G.J. Jacobi, Ueber Gauss neue Methode, die Werthe der Integrale näherungsweise zu finden, J. Reine Angew. Math. 1 (1826) 301–308. Reprinted in: Gesammelte Werke, vol. 6, Reimer, Berlin, 1891, pp. 3–11.

[55] C.G.J. Jacobi, Über die Reduction der quadratischen Formen auf die kleinste Anzahl Glieder, J. Reine Angew. Math. 39 (1850) 290–292. Reprinted in: Gesammelte Werke, vol. 6, Reimer, Berlin, 1891, pp. 318–320.

[56] J. Kleinberg. Authoritative sources in a hyperlinked environment, J. ACM 46 (1999) 604–632.

[57] G. Kollias, E. Gallopoulos, Functional rankings with multidamping: Generalizing pagerank with inhomogeneous matrix products, Tech. Rep. TR HPCLAB-SCG 01/09-11, University of Patras, Greece, 2011.

[58] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, J. Research Nat. Bur. Standards 45 (1950) 225–282.

[59] C. Lanczos, Solution of systems of linear equations by minimized iterations, J. Research Nat. Bur. Standards 49 (1952) 33–53.

[60] A.N. Langville, C.D. Meyer, A survey of eigenvector methods for Web information retrieval, SIAM Rev. 47 (2005) 135–161.

[61] A.N. Langville, C.D. Meyer, Google's pagerank and beyond: The science of search engine rankings, Princeton University Press, Princeton, NJ, 2006.

[62] A.N. Langville, C.D. Meyer, Who's #1? The science of rating and ranking, Princeton University Press, Princeton, NJ, 2012.

[63] J. Liesen, Z. Strakoš, Krylov subspace methods: Principles and analysis, Oxford University Press, Oxford, 2013.

[64] F. Marcelán, R. Álvarez-Nodarse, On the "Favard theorem" and its extension, J. Comput. Appl. Math. 127 (2001) 231–254.

[65] J.R. Magnus, H. Neudecker, Matrix differential calculus with applications in statistics and econometrics, J. Wiley & Sons Inc., Chichester, 1988.

[66] G. Meurant, Z. Strakoš, The Lanczos and conjugate gradient algorithms in finite precision arithmetic, Acta Numer. 15 (2006) 471–542.

[67] M.E.J. Newman, The structure and function of complex networks, SIAM Rev. 45 (2003) 167–256.

[68] M.E.J. Newman, Networks: An introduction, Cambridge University Press, Cambridge, UK, 2010.

[69] M.E.J. Newman, A.-L. Barabasi, D.J. Watts, The structure and dynamics of networks, Princeton University Press, Princeton, NJ, 2003.

[70] B.N. Parlett, The Symmetric eigenvalue problem, Prentice-Hall, Englewood Cliffs, 1980.

[71] B.N. Parlett, Reduction to tridiagonal form and minimal realizations, SIAM J. Matrix Anal. Appl. 13 (1992) 567–593.

[72] B.N. Parlett, D.R. Taylor, Z.A. Liu, A look-ahead Lanczos algorithm for unsymmetric matrices, Math. Comp. 44 (1985) 105–124.

[73] T. Popoviciu, Sur une généralisation de la formule d'integration numérique de Gauss, Acad. R. P. Romîne Fil. Iaşi Stud. Cerc. Şti. 6 (1955) 29–57.

[74] S. Pozza, M. Pranić, Z. Strakoš, Gauss quadrature for quasi-definite linear functionals, submitted.

[75] H. Richter, Über Matrixfunktionen, Math. Ann. 122 (1950) 16–34.

[76] R.F. Rinehart, The equivalence of definitions of a matrix function, Amer. Math. Monthly 62 (1955) 395–414.

[77] Y. Saad, Iterative methods for sparse linear systems, SIAM, Philadelphia, 2003.

[78] P.E. Saylor, D.C. Smolarski, Why Gauss quadrature in the complex plane?, Numer. Algorithms 26 (2001) 251–280.

[79] H. Schwerdtfeger, Les fonctions de matrices I. Les fonctions univalentes, in: Actualités Scientifiques et Industrielles No. 649, Herman, Paris, 1938.

[80] N.H. Scott, A Theorem on isotropic null vectors and its application to thermoelasticity, Proceedings: Mathematical and Physical Sciences 440 (1993) 431–442.

[81] B. Simon, CMV matrices: five years after, J. Comput. Appl. Math. 208 (2007) 120–154.

[82] A. Spitzbart, A Generalization of Hermite's interpolation formula, Amer. Math. Monthly 67(1) (1960) 42–46.

[83] D.D. Stancu, A.H. Stroud, Quadrature formulas with simple gaussian nodes and multiple fixed nodes, Math. Comput. 17(84) (1963) 384–394.

[84] T.J. Stieltjes, Recherches sur les fractions continues, Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys. 8 (1894) J. 1–122. Reprinted in: Oeuvres, vol. 2, P. Noordhoff, Groningen, 1918, pp. 402–566. English translation, Investigation on continued fractions, in: Thomas Jan Stieltjes, Collected Papers, vol. 2, Springer-Verlag, Berlin, 1993, pp. 609–745.

[85] Z. Strakoš, Model reduction using the Vorobyev moment problem, Numer. Algorithms 51 (2009) 363–379.

[86] J.J. Sylvester, On the equation to the secular inequalities in the planetary theory, Phil. Mag. (5)16 (1883) 267–269.

[87] G. Szegö, Orthogonal polynomials, Amer. Math. Soc. Colloquium Publications, 23, New York, 1939.

[88] P. Tsaparas, Link analysis ranking, Ph.D. Thesis, Department of Computer Science, University of Toronto, 2004.

[89] A. van der Sluis, H.A. van der Vorst, The convergence behavior of Ritz values in the presence of close eigenvalues, Linear Algebra Appl. 88/89 (1987) 651–694.

[90] S. Vigna, Spectral ranking, arXiv:0912.0238v13 [cs.IR], November 8, 2013.

[91] Y.V. Vorobyev, Methods of moments in applied mathematics, Translated from the Russian by Bernard Seckler, Gordon and Breach Science Publishers, New York, 1965.

[92] H.S. Wall, Analytic theory of continued fractions, Chelsea Pub. Co., Bronx, N.y., 1948.

[93] E. Weyr, Note sur la théorie de quantités complexes formées avec n unités principales, Bull. Sci. Math. II 11 (1887) 205–215.

[94] H. Wilf, Mathematics for the physical sciences, J. Wiley & Sons Inc., London, 1962.

[95] J.H. Wilkinson, The algebraic eigenvalue problem, Clarendon Press, Oxford, 1988.

Part II

Sequence Transformations

CHAPTER 1

Sequence Transformations

In many occasions, in numerical analysis we deal with sequences slowly converging to their limits. Frequently they converge so slowly that it becomes impractical to effectively use them. For this reason sequence transformations have a fundamental role since they could potentially accelerate the convergence of a sequence.

Sequence transformations are a really useful numerical tool, and a vast literature has been developed. We refer to, e.g., [7], [11], [12], [15], [28], [32], [37]. In addition, in these last years many works have appeared on how to effectively use sequence transformations in practical situations. See for example [5], [24], [26], and [30]. We refer also to [19], in which is discussed the efficiency of many numerical techniques for the evaluation of power series expansions for special functions.

1.1 Introduction

In this chapter we give an introduction to convergence acceleration. In particular, we follow and refer to the book by Brezinski and Redivo-Zaglia [12].

Let (S_n) be a sequence of real or complex numbers converging to a finite number S. We want to define a transformation T from a set of

sequences to another one, i.e., $T : (S_n) \mapsto (T_n)$, with (T_n) a sequence with the following properties

1. (T_n) converges;
2. (T_n) converges to the same limit as (S_n), i.e., S;
3. (T_n) converges *faster* than (S_n), i.e., $\lim_{n\to\infty}(T_n - S)/(S_n - S) = 0$.

If the sequence transformation T gives the sequence T_n satisfying only properties 1. and 2. we say that T is *regular* for the sequence (S_n). While, if sequences T_n satisfy property 3 we say that T *accelerates* the convergence of (S_n).

Delahaye and Germain-Bonne in [16, 15] proved that an universal transformation able to accelerate any converging sequence does not exist. In particular, in [15, 17, 21, 22] it was proved that it is impossible to give a transformation able to accelerate all the sequences in some sets, in particular:

- the set of *monotone* sequences, i.e., sequences such that $S_{n+1} \geq S_n$ or $S_{n+1} \leq S_n$ for $n = 0, 1, \ldots$;
- the set of *logarithmic* sequences, i.e., sequences such that $\lim_{n\to\infty}(S_{n+1} - S)/(S_n - S) = 1$.

Then, in practical situations it is important to develop specific algorithms for the class of sequences of interest. However, if this class is too small, such a transformation will be useful only in particular cases; on the other hand, a specialization typically provides a faster acceleration.

A sequence transformation T can be represented by infinite sets of doubly indexed quantities T_n^k, for $n, k = 0, 1, \ldots$. Typically, n is the minimal index of the sequence elements $S_n, \ldots, S_{n+\ell}$ which are used for the computation of the transformation T_n^k, and k is a measure for the complexity of the computation of T_n^k, as for example the number of sequence elements necessary to compute it. Usually, the transformed sequence is given by an index-constant path, i.e., the sequence $T_n^k, T_{n+1}^k, T_{n+2}^k, \ldots$ with fixed minimal index k and $n \to \infty$. However, in other cases, the transformation order n is kept fixed and $k \to \infty$, i.e., we get an order-constant path $T_n^k, T_n^{k+1}, T_n^{k+2}, \ldots$. For details and a further discussion about different paths we refer to [33, Section 2]. In principle, the index-constant approach is more efficient since it uses more available input data. However, we will consider and define order-constant transformations.

Let $S_n, \ldots, S_{n+\ell}$ the elements needed to compute T_n, the n-th elements of the transformed sequence obtained applying T to the sequence (S_n). Then for some sequence it is possible that

$$\lim_{n\to\infty} \frac{T_n - S}{S_n - S} = 0, \qquad \text{while} \qquad \lim_{n\to\infty} \frac{T_n - S}{S_{n+j} - S} \neq 0,$$

for some j between $1, \ldots, \ell$. This means that T_n is not faster than the sequence given by (S_{n+j}). Hence, it would be better to study the convergence rate of a sequence transformation looking at the ratio $(T_n - S)/(S_{n+\ell} - S)$. However, notice that

$$\frac{T_n - S}{S_{n+\ell} - S} = \frac{T_n - S}{S_n - S} \frac{S_n - S}{S_{n+1} - S} \cdots \frac{S_{n+\ell-1} - S}{S_{n+\ell} - S}.$$

Hence, if $(S_{n+1} - S)/(S_n - S) \neq 0$ for every n, and does not tend to zero, then $\lim_{n\to\infty}(T_n - S)/(S_n - S) = 0$ if and only if $\lim_{n\to\infty}(T_n - S)/(S_{n+\ell} - S) = 0$. However, if

$$\lim_{n\to\infty} \frac{S_{n+1} - S}{S_n - S} = 0$$

we say that the sequence (S_n) is *hyperlinearly* convergent and, in practice, we can exclude this case from our analysis since it does not need to be accelerated.

In the analysis of a sequence transformation, the notion of *kernel* is particularly useful. We define the kernel of a transformation $T : (S_n) \longmapsto (T_n)$ as the set of all sequences (S_n) which are transformed by T into a constant sequence, i.e., for every sequence in the kernel there exists S for which $T_n = S$ for all n, or eventually for every $n \geq N$, with $N > 0$. Usually, S is the limit of the sequence, if it exists. The importance of the kernel came from the fact that, even if it has not yet been proven, the "closer" a sequence is to the kernel, the faster the transformed sequence converges to the same limit, as numerical experiments have always confirmed.

The standard way of defining a transformation is to start from the kernel. In particular, we can express it by an implicit relation R that consider ℓ elements of a sequence and a value S, i.e.,

$$R(S_n, \ldots, S_{n+\ell}, S) = 0.$$

We say that a sequence (S_n) is in the kernel $\mathcal{K}$ if it satisfies the previous equation for every n. Moreover, we call *extrapolation method* every

sequence transformation $T : (S_n) \longmapsto (T_n)$ for which $T_n = S$ for every n, if $(S_n) \in \mathcal{K}$. The name *extrapolation* comes from *interpolation* and it is explained by the procedure to build a transformation from its kernel.

Let $S_n, \ldots, S_{n+\ell+m}$ be given, and (u_n) a sequence in the kernel $\mathcal{K}$ which satisfies the interpolation conditions

$$u_i = S_i, \quad \text{for } i = n, \ldots, n+\ell+m.$$

Since, (u_n) is in the kernel, it satisfies the implicit conditions

$$R(u_i, \ldots, u_{i+\ell}, S) = 0, \quad \text{for } i = 1, 2, \ldots,$$

which we assume depends on m parameters $a_1, \ldots, a_m$. Then, using the interpolation conditions we get

$$R(S_i, \ldots, S_{i+\ell}, S) = 0, \quad \text{for } i = n, \ldots, m,$$

that is a system of $m+1$ equations in $m+1$ unknowns $a_1, \ldots, a_m, S$. Hence, if we solve this system we obtain S. Notice that the computed value of S depend on n and $k = \ell + m$. Then, we can define the transformation setting $T_n^k = S$. Notice that if R is linear with respect to the unknowns $a_1, \ldots, a_m, S$, then T_n^k can be expressed as the solution of a linear system, and, hence, as a ratio of two determinants.

Now, we will obtain the well-known Aitken's Δ^2 process starting from its kernel. Doing a short digression we recall that the method is named after Aitken since Alexander Craig Aitken (1895-1967) used it in [3] (1926). However, the Aitken's Δ^2 process was actually discovered by Japanese Mathematician Takakazu Seki (?-1708) before 1680. The same method was obtained by Hans von Naegelsbach (1838-?) in 1876 and by James Clerk Maxwell (1831-1879) in 1873 but none of them used it for the purpose of acceleration; see, e.g., [9, 25].

Let us consider the kernel given by sequences of the form

$$S_n = S + a\lambda^n, \quad n = 0, 1, \ldots, \tag{1.1}$$

where $a \in \mathbb{C}$ is different from 0 and $\lambda \in \mathbb{C}$ is different from 0 and 1. If $|\lambda| < 1$, then S is the limit of the sequence; otherwise, for (S_n) diverging, S is called the *antilimit* of the sequence. Then, the implicit form of the kernel is

$$u_{i+1} - S = \lambda(u_i - S), \quad \text{for } i = 1, 2, \ldots$$

or equivalently

$$R(u_i, u_{i+1}, S) = a_1(u_i - S) + a_2(u_{i+1} - S) = 0 \quad \text{for } i = 1, 2, \ldots, \tag{1.2}$$

with $a_1, a_2 \neq 0$. Then, given the values S_n, S_{n+1} we get the system

$$\begin{cases} a_1(S_n - S) + a_2(S_{n+1} - S) = 0 \\ a_1(S_{n+1} - S) + a_2(S_{n+2} - S) = 0 \end{cases}$$

Solving it and setting $T_n = S$ we obtain

$$T_n = \frac{S_n S_{n+2} - S_{n+1}^2}{S_{n+2} - 2S_{n+1} + S_n}, \quad n = 0, 1, \dots ,$$

which is the Aitken's Δ^2 process.

Notice that, by construction, we have proved that the kernel of Aitken's Δ^2 process consists of all the sequences of the form of (1.1) and only them. However, sufficiency is usually difficult to prove for a general sequence transformation.

The previous formula is unstable since, when S_n, S_{n+1}, S_{n+2} are almost equal, cancellation errors arise in the denominator and in the numerator; see, e.g., [12, p.34 - 35, pp.400 - 403],[13, p. 173]. Then we can give the following more stable equivalent formulas

$$T_n = S_n - \frac{(\Delta S_n)^2}{\Delta^2 S_n} = S_{n+1} - \frac{\Delta S_n \Delta S_{n+1}}{\Delta^2 S_n} = S_{n+2} - \frac{(\Delta S_{n+1})^2}{\Delta^2 S_n}, \quad (1.3)$$

for $n = 0, 1, \dots$, where Δ is the forward difference operator defined as

$$\Delta S_n = S_{n+1} - S_n.$$

Notice that in all the formulas the denominator is $\Delta^2 S_n = \Delta S_{n+1} - \Delta S_n$, which explains the name of the method. For more details on Aitken's Δ^2 process, we refer to [12, Chapter 1].

In the following sections we will present the definitions of some well-known sequence transformations we will use in the numerical experiments of Chapter 2: Shanks' transformation and ε-algorithm, θ-algorithm, and Levin's algorithm.

1.2 Shanks' Transformation

Let us generalize the Kernel of Aitken's process (1.2) considering the kernel given by the implicit relation

$$R(u_i, \dots, u_{i+\ell}, S) = a_1(u_i - S) + \cdots + a_{\ell+1}(u_{i+\ell} - S) = 0 \quad (1.4)$$

with $a_{\ell+1} \neq 0$ and $a_1 + \cdots + a_\ell \neq 0$, for $i = 1, 2, \ldots$. We now follow the procedure shown in the previous chapter, setting $m = \ell$, $u_i = S_i$ for $i = n, \ldots, n + 2\ell$, and $T_n = S$. Moreover, since $R(u_i, \ldots, u_{i+\ell}, S) = 0$ is invariant for multiplication by scalars different from zero, we can assume $a_1 + \cdots + a_{\ell+1} = 1$. Then we get the following linear system

$$\begin{cases} S_n &= T_n + b_1 \Delta S_n + \cdots + b_\ell \Delta S_{n+\ell-1} \\ S_{n+1} &= T_n + b_1 \Delta S_{n+1} + \cdots + b_\ell \Delta S_{n+\ell} \\ \vdots & \qquad\qquad \vdots \\ S_{n+\ell} &= T_n + b_1 \Delta S_{n+\ell} + \cdots + b_\ell \Delta S_{n+2\ell-1}, \end{cases}$$

with $T_n = S$, $a_i = b_i - b_{i-1}$, for $i = 2, \ldots, \ell + 1$, $a_1 = b_1 + 1$ and $b_{\ell+1} = 0$. Using the classical determinant formula for the solution of a linear system we get

$$T_n = \frac{\begin{vmatrix} S_n & S_{n+1} & \ldots & S_{n+\ell} \\ \Delta S_n & \Delta S_{n+1} & \ldots & \Delta S_{n+\ell} \\ \vdots & \vdots & & \vdots \\ \Delta S_{n+\ell-1} & \Delta S_{n+\ell} & \ldots & \Delta S_{n+2\ell-1} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \ldots & 1 \\ \Delta S_n & \Delta S_{n+1} & \ldots & \Delta S_{n+\ell} \\ \vdots & \vdots & & \vdots \\ \Delta S_{n+\ell-1} & \Delta S_{n+\ell} & \ldots & \Delta S_{n+2\ell-1} \end{vmatrix}}.$$

This transformation is known as *Shanks' transformation* and it is usually denoted as $e_\ell(S_n)$. It was introduced by Shanks in [27]. In [10] Brezinski and Crouzeix proved that every sequence in the Kernel of the Shanks' transformation can be written in the form (1.4). In addition, they show that these sequences can be explicitly written as

$$S_n = S + \sum_{i=1}^{m} A_i(n) r_i^n + \sum_{i=m+1}^{k} \left[B_i(n) \cos(b_i n) + C_i(n) \sin(b_i n) \right] e^{w_i n} + \\ + \sum_{i=0}^{t} c_i \delta_{in},$$

where $r_i \neq 1$ for $i = 1, \ldots, m$, δ_{in} is the Kronecker's symbol, and A_i, B_i, C_i are polynomials in n such that

$$t + \sum_{i=1}^{m} d_i + 2 \sum_{i=m+1}^{k} d_i = \ell - 1,$$

with d_i the degree of A_i plus one for $i = 1, \ldots, m$ and the maximum of the degrees between B_i and C_i plus one for $i = m+1, \ldots, k$. We use the convention $t = -1$ if the last sum in the formula does not appear.

The most important method for the computation of $e_k(S_n)$ is the so called *ε-algorithm* introduced by Wynn in [38]. It consists in computing scalars $\varepsilon_k^{(n)}$ following the rules

$$\varepsilon_{-1}^{(n)} = 0, \quad \varepsilon_0^{(n)} = S_n, \quad n = 0, 1, \ldots \tag{1.5}$$

$$\varepsilon_{k+1}^{(n)} = \varepsilon_{k-1}^{(n+1)} + \frac{1}{\varepsilon_k^{(n+1)} - \varepsilon_k^{(n)}}, \quad k, n = 0, 1, \ldots. \tag{1.6}$$

The ε-algorithm is related to Shanks' transformation by the equivalence

$$\varepsilon_{2k}^{(n)} = e_k(S_n), \quad \text{for } k, n = 0, 1, \ldots,$$

which was proved by Beckermann in [4]. The values of $\varepsilon_k^{(n)}$ can be represented in the following double entry table in which k is the index of the columns

$$\begin{array}{llllll}
\varepsilon_{-1}^{(0)} = 0 & & & & & \\
& \varepsilon_0^{(0)} = S_0 & & & & \\
\varepsilon_{-1}^{(1)} = 0 & & \varepsilon_1^{(0)} & & & \\
& \varepsilon_0^{(1)} = S_1 & & \varepsilon_2^{(0)} & & \\
\varepsilon_{-1}^{(2)} = 0 & & \varepsilon_1^{(1)} & & \varepsilon_3^{(0)} & \\
& \varepsilon_0^{(2)} = S_2 & & \varepsilon_2^{(1)} & & \ddots \\
\varepsilon_{-1}^{(3)} = 0 & & \varepsilon_1^{(2)} & & \varepsilon_3^{(1)} & \\
\vdots & \vdots & & \ddots & & \ddots
\end{array}$$

We remark that the rules of the ε-algorithm relates the quantities at the four corners of a rhombus in the table

$$\begin{array}{ccccc}
 & & \varepsilon_k^{(n)} & & \\
 & \nearrow & & \searrow & \\
\varepsilon_{k-1}^{(n+1)} & & & & \varepsilon_{k+1}^{(n)} \\
 & \searrow & & \nearrow & \\
 & & \varepsilon_k^{(n+1)} & &
\end{array}.$$

Hence, following this scheme, we can compute all the table elements over the diagonal from $\varepsilon_0^{(n)} = S_n$ to $\varepsilon_n^{(0)}$ knowing $S_0, \ldots, S_n$ values.

1.3 θ-algorithm

The θ-algorithm was first proposed by Brezinski in [6]. We will describe it following Section 2.9 of [12]. First, notice that we can rewrite the rule (1.6) of the ε-algorithm as follow

$$\varepsilon_{k+1}^{(n)} = \varepsilon_{k-1}^{(n+1)} + D_k^{(n)},$$

with $D_k^{(n)} = \left(\varepsilon_k^{(n+1)} - \varepsilon_k^{(n)}\right)^{-1}$.

Before going on we recall the following theorem.

Theorem 1.1 ([12, Theorem 1.25])**.** *Assuming tha there exist λ, ρ such that*

$$\lim_{n\to\infty} \frac{\Delta T_{n+1}}{\Delta T_n} = \rho, \qquad \text{and} \qquad \lim_{n\to\infty} \frac{\Delta S_{n+1}}{\Delta S_n} = \lambda,$$

with $|\lambda| < 1$ and $|\rho| < 1$. Then,

$$\lim_{n\to\infty} \frac{T_n - S}{S_n - S} = a$$

if and only if

$$\lim_{n\to\infty} \frac{\Delta T_n}{\Delta S_n} = a,$$

with $a \in \mathbb{C}$.

For a proof we refer to [8, Theorem 5] and [31, Theorem 3.5].

If sequences $\varepsilon_{2k+2}^{(n)}$ and $\varepsilon_{2k}^{(n)}$ satisfy the assumptions of the previous theorem, then

$$\lim_{n\to\infty} \frac{\Delta\varepsilon_{2k+2}^{(n)}}{\Delta\varepsilon_{2k}^{(n)}} = 0$$

implies

$$\lim_{n\to\infty} \frac{\varepsilon_{2k+2}^{(n)} - S}{\varepsilon_{2k}^{(n)} - S} = 0.$$

It means that $\varepsilon_{2k+2}^{(n)}$ converges faster than $\varepsilon_{2k}^{(n)}$ if and only if

$$\lim_{n\to\infty} \frac{\Delta D_{2k+1}^{(n)}}{\Delta\varepsilon_{2k}^{(n+1)}} = -1. \tag{1.7}$$

Hence, when this last condition is not satisfied we can introduce a parameter ω_k in the algorithm obtaining the new rule

$$\varepsilon_{2k+2}^{(n)} = \varepsilon_{2k}^{(n+1)} + \omega_k D_{2k+1}^{(n)}.$$

Then, if we set

$$\omega_k = -\lim_{n\to\infty} \frac{\Delta\varepsilon_{2k}^{(n+1)}}{\Delta D_{2k+1}^{(n)}}$$

the new sequence satisfies condition 1.7. From a practical point of view the computation of ω_k is difficult since it involves a limit. Hence, we will give it as

$$\omega_k = -\frac{\Delta\varepsilon_{2k}^{(n+1)}}{\Delta D_{2k+1}^{(n)}}.$$

We can give the rule for the θ-algorithm.

$$\begin{aligned}
\theta_{-1}^{(n)} = 0, \quad \theta_0^{(n)} = S_n, \quad & n = 0, 1, \dots \\
\theta_{2k+1}^{(n)} = \theta_{2k-1}^{(n+1)} + D_{2k}^{(n)}, \quad & k, n = 0, 1, \dots \\
\theta_{2k+2}^{(n)} = \theta_{2k}^{(n+1)} - \frac{\Delta\theta_{2k}^{(n+1)}}{D_{2k+1}^{(n)}} D_{2k+1}^{(n)}, \quad & k, n = 0, 1, \dots,
\end{aligned}$$

with $D_k^{(n)} = \left(\theta_k^{(n+1)} - \theta_k^{(n)}\right)^{-1}$.

Finally, we recall the following result about the kernel of θ-algorithm for $k = 2$.

Theorem 1.2 ([12, Theorem 2.36])**.** *A sequence (S_n) is in the kernel of θ-algorithm with $k = 2$ (i.e., $\theta_2^{(n)} = S$ for $n = 0, 1, \dots$) if and only if the sequence has one of the following form*

1. *$S_n = S + (S_0 - S)\lambda^n$, with $S_0 \neq S$ and $\lambda \neq 0, 1$;*
2. *$S_n = S + (S_0 - S)\prod_{i=0}^{n-1}[1 - d(i-m)^{-1}]$, with $S_0 \neq S$, $d \neq 1$ and $m, m + d$ not integers;*
3. *$S_0 = S$ and $S_n = S + (S_1 - S)\prod_{i=0}^{n-1}(1 - di^{-1})$ for $n = 1, 2, \dots$, with $S_1 \neq S$ and d not an integer.*

1.4 Levin type Transformation

Levin's transformation was introduced by Levin in [23] and can be considered a generalization of the Aitken's Δ^2 process. In order to obtain

it we will follow the path described in [12, Section 2.7]. As we have seen in Section 1.1 the sequences in the kernel of the Aitken's Δ^2 process are of the form

$$S_n - S = a\,\Delta S_n, \quad n = 0, 1, \ldots.$$

We replace the constant a with a polynomial of degree $k-1$ in the variable $(n+b)^{-1}$, with b a non-zero real coefficient different from the negative integers. Moreover, instead of ΔS_n we introduce an auxiliary sequence $(g(n))$. Hence, we consider sequences of the kind

$$S_n - S = g(n)\,\left(a_1 + a_2(n+b)^{-1} + \cdots + a_k(n+b)^{-(k-1)}\right), \quad n = 0, 1, \ldots.$$

Multiplying both sides of this equation by $(n+b)^{k-1}$ gives

$$(n+b)^{k-1}\frac{S_n - S}{g(n)} = a_1(n+b)^{k-1} + a_2(n+b)^{k-2} + \cdots + a_k, \quad n = 0, 1, \ldots.$$

Let Δ^k the operator obtained applying k times the operator Δ. Then $\Delta^k p(n) = 0$ for every polynomial p of degree $k-1$. Hence, if we apply Δ^k to the previous equation we get

$$\Delta^k\left((n+b)^{k-1}\frac{S_n - S}{g(n)}\right) = 0,$$

for $n = 0, 1, \ldots$. Moreover, since Δ^k is a linear operator we have

$$\Delta^k\left((n+b)^{k-1}\frac{S_n}{g(n)}\right) = S\,\Delta^k\left(\frac{(n+b)^{k-1}}{g(n)}\right).$$

Setting $L_k^{(n)} = S$, then we get the transformation

$$L_k^{(n)} = \frac{\Delta^k\left(S_n\,(n+b)^{k-1}/g(n)\right)}{\Delta^k\left((n+b)^{k-1}/g(n)\right)}, \quad n = 0, 1, \ldots.$$

With the following choices for $g(n)$ we obtain the Levin transformations:

- *u-transform*: $g(n) = (n+b)\Delta S_{n-1}$;
- *t-transform*: $g(n) = \Delta S_{n-1}$;
- *v-transform*: $g(n) = -\Delta S_{n-1}\Delta S_n/\Delta^2 S_{n-1}$.

Finally, we remark that Levin's transformation can be computed by a recursive algorithm, see [18].

CHAPTER 2

Generalizations of Aitken's Process

In this chapter we present the results we obtained in [14]. In [13], Brezinski and Redivo-Zaglia considered two kernels consisting of sequences of the form

$$S_n = S + (a + bx_n)\lambda^n, \quad n = 0, 1, \ldots,$$

or

$$S_n = S + (a + bx_n)^{-1}\lambda^n, \quad n = 0, 1, \ldots,$$

where S, a, b and λ are unknown numbers and (x_n) a known sequence. These kernels obviously contain Aitken's Δ^2 process kernel (1.1).

In this chapter, we will construct several sequence transformations whose kernel is another generalization of kernel (1.1), consisting of sequences of the form

$$S_n = S + a_n\lambda^n, \quad n = 0, 1, \ldots,$$

where S and λ are unknown parameters, and (a_n) is a known sequence.

After defining them (Section 2.1), we will give some results on their acceleration properties (Section 2.2). Finally, in Section 2.3, we will present numerical tests on the best of this transformations and we will compare it with other well-known transformations.

2.1 New Transformations

We consider a kernel containing sequences of the form

$$S_n = S + a_n\lambda^n, \quad n = 0, 1, \ldots, \tag{2.1}$$

with S and λ unknown parameters, and (a_n) a known sequence.

We remark that the convergence of sequences of the form (2.1) depends on the term $a_n\lambda^n$ as $n \to \infty$. Therefore, both (a_n) and λ determine the convergence of the sequence. For example, if $a_n \equiv a$, then the sequence converges if and only if $|\lambda| \leq 1$. We remark that if $\lambda = 1$ then the limit of the sequence is $S + a$, with $a = \lim_{n\to\infty} a_n$. In addition, for a slowly increasing (a_n) we will have convergence only when $|\lambda| < 1$. We will make further comments on convergence in Section 2.2.

First, we will introduce sequence transformations with kernel containing the sequences of the kind (2.1), where (a_n) is a given sequence and S, λ are unknowns. As done in Section 1.1, we will express S as a function using the kernel, i.e.

$$S = f(S_n, \ldots, S_{n+k}; a_n, \ldots, a_{n+\ell})$$

for $n = 0, 1, \ldots$ and $k, \ell \in \mathbb{N}$, in order to compute $a_n, \ldots, a_{n+\ell}$ using an interpolation process depending on n. Then we can define the transformation as follows

$$T_n := f(S_n, \ldots, S_{n+k}; a_n, \ldots, a_{n+\ell}), \quad \text{for } n = 0, 1, \ldots.$$

The first transformation ${}^1\mathrm{T}_n$ is given by solving a linear system. The other two transformations are similar, but each of them needs a different estimate of the parameter λ.

Considering the kernel (2.1) for indexes n and $n + 1$ gives

$$\begin{array}{rcl} S_n & = & S + a_n\lambda^n \\ S_{n+1} & = & S + a_{n+1}\lambda^{n+1}. \end{array} \tag{2.2}$$

By the first equation we get $\lambda^n = (S_n - S)/a_n$. Therefore, the second equation can be rewritten as

$$a_nS - a_{n+1}\lambda S + a_{n+1}S_n\lambda = a_nS_{n+1}. \tag{2.3}$$

Since (2.3) is nonlinear with respect to S and λ, we consider the linear system with unknowns $S, \lambda S$ and λ obtained by (2.3) for indexes $n, n+1, n+2$

$$\begin{aligned} a_n S - a_{n+1}\lambda S + a_{n+1}S_n\lambda &= a_n S_{n+1}, \\ a_{n+1} S - a_{n+2}\lambda S + a_{n+2}S_{n+1}\lambda &= a_{n+1} S_{n+2}, \\ a_{n+2} S - a_{n+3}\lambda S + a_{n+3}S_{n+2}\lambda &= a_{n+2} S_{n+3}. \end{aligned} \tag{2.4}$$

If we compute S by the system (2.4) as a ratio of determinants we can define the first new transformation

$$\begin{aligned} {}^1\mathrm{T}_n &= \frac{\begin{vmatrix} a_n S_{n+1} & -a_{n+1} & a_{n+1}S_n \\ a_{n+1}S_{n+2} & -a_{n+2} & a_{n+2}S_{n+1} \\ a_{n+2}S_{n+3} & -a_{n+3} & a_{n+3}S_{n+2} \end{vmatrix}}{\begin{vmatrix} a_n & -a_{n+1} & a_{n+1}S_n \\ a_{n+1} & -a_{n+2} & a_{n+2}S_{n+1} \\ a_{n+2} & -a_{n+3} & a_{n+3}S_{n+2} \end{vmatrix}} \qquad (2.5) \\ &= \frac{\begin{vmatrix} a_{n+1} & a_{n+2} & a_{n+3} \\ a_n S_{n+1} & a_{n+1}S_{n+2} & a_{n+2}S_{n+3} \\ a_{n+1}S_n & a_{n+2}S_{n+1} & a_{n+3}S_{n+2} \end{vmatrix}}{\begin{vmatrix} a_{n+1} & a_{n+2} & a_{n+3} \\ a_n & a_{n+1} & a_{n+2} \\ a_{n+1}S_n & a_{n+2}S_{n+1} & a_{n+3}S_{n+2} \end{vmatrix}} = \frac{N_n}{D_n}. \qquad (2.6) \end{aligned}$$

The numerator N_n and the denominator D_n can be written as

$$\begin{aligned} N_n &= a_{n+3}\Delta S_{n+1}(a_{n+1}^2 S_{n+2} - a_n a_{n+2} S_{n+1}) \\ &\quad - a_{n+1}\Delta S_n (a_{n+2}^2 S_{n+3} - a_{n+1}a_{n+3}S_{n+2}), \\ D_n &= a_{n+3}\Delta S_{n+1}(a_{n+1}^2 - a_n a_{n+2}) - a_{n+1}\Delta S_n (a_{n+2}^2 - a_{n+1}a_{n+3}) \\ &= a_{n+3}S_{n+2}(a_{n+1}^2 - a_n a_{n+2}) + a_{n+2}S_{n+1}(a_n a_{n+3} - a_{n+1}a_{n+2}) \\ &\quad + a_{n+1}S_n (a_{n+2}^2 - a_{n+1}a_{n+3}). \end{aligned}$$

As it is well-known, a transformation expressed in the previous way is unstable. However, ${}^1\mathrm{T}_n$ can be rewritten in the equivalent form, ${}^1\mathrm{T}_n = S_{n+i} - (S_{n+i}D_n - N_n)/D_n$, for $i = 0, 1, 2, 3$. Simplifying the numerator

gives the following alternative expressions

$$
\begin{aligned}
{}^{1}\mathrm{T}_n &= S_n - \frac{1}{D_n}\Big[-a_{n+1}^2 a_{n+3}\left(\Delta S_{n+1}+\Delta S_n\right)^2 \\
&\quad + \left(a_n a_{n+2} a_{n+3}\Delta S_{n+1} + a_{n+1}a_{n+2}^2\left(S_{n+3}-S_n\right)\right)\Delta S_n\Big] \\
&= S_{n+1} - \frac{1}{D_n}\Big[a_{n+1}a_{n+2}^2\Delta S_n\left(\Delta S_{n+2}+\Delta S_{n+1}\right) \\
&\quad - a_{n+1}^2 a_{n+3}\Delta S_{n+1}\left(\Delta S_{n+1}+\Delta S_n\right)\Big] \\
&= S_{n+2} - \frac{1}{D_n}\Big[a_{n+1}a_{n+2}^2\Delta S_n\Delta S_{n+2} \\
&\quad - a_n a_{n+2} a_{n+3}\left(\Delta S_{n+1}\right)^2\Big] \\
&= S_{n+3} - \frac{1}{D_n}\Big[a_{n+1}^2 a_{n+3}\Delta S_{n+2}\left(\Delta S_{n+1}+\Delta S_n\right) \\
&\quad - a_n a_{n+2} a_{n+3}\Delta S_{n+1}\left(\Delta S_{n+1}+\Delta S_{n+2}\right)\Big].
\end{aligned}
\tag{2.7}
$$

If we assume $a_n \neq 0$ for all $n \in \mathbb{N}_0$ we can give another equivalent expression of the transformation. Let us divide the i-th column of both the numerator and the denominator of (2.5) by a_{n+i} and replace the second and third column by their difference with the preceding ones. Then, the determinants have only two rows and columns. Setting $\beta_n = a_n/a_{n+1}$ gives

$$
{}^{1}\mathrm{T}_n = \frac{\begin{vmatrix} \Delta(\beta_n S_{n+1}) & \Delta(\beta_{n+1}S_{n+2}) \\ \Delta S_n & \Delta S_{n+1} \end{vmatrix}}{\begin{vmatrix} \Delta\beta_n & \Delta\beta_{n+1} \\ \Delta S_n & \Delta S_{n+1} \end{vmatrix}}.
$$

Assuming $\Delta S_n \neq 0$ for $n = 0, 1, \ldots$ we can divide the i-th column by ΔS_{n+i-1}, for $i = 1, 2$, obtaining a compact form of the transformation ${}^{1}\mathrm{T}_n$

$$
{}^{1}\mathrm{T}_n = \frac{\Delta\left(\dfrac{\Delta\left(\beta_n S_{n+1}\right)}{\Delta S_n}\right)}{\Delta\left(\dfrac{\Delta\beta_n}{\Delta S_n}\right)}. \tag{2.8}
$$

Similarly to (2.7) we get some equivalent formulations

$$\begin{aligned}
{}^1\mathrm{T}_n &= S_n + \frac{\Delta^2\beta_n + \Delta\beta_{n+1}\dfrac{\Delta S_n}{\Delta S_{n+1}} + \Delta\left(\beta_{n+1}\dfrac{\Delta S_{n+1}}{\Delta S_n}\right)}{\Delta\left(\dfrac{\Delta\beta_n}{\Delta S_n}\right)} \\
&= S_{n+1} + \frac{\Delta\beta_{n+1} + \Delta\left(\beta_{n+1}\dfrac{\Delta S_{n+1}}{\Delta S_n}\right)}{\Delta\left(\dfrac{\Delta\beta_n}{\Delta S_n}\right)} \\
&= S_{n+2} + \frac{\Delta\beta_{n+1}\dfrac{\Delta S_{n+2}}{\Delta S_{n+1}} + \Delta\left(\beta_n\dfrac{\Delta S_{n+1}}{\Delta S_n}\right)}{\Delta\left(\dfrac{\Delta\beta_n}{\Delta S_n}\right)} \\
&= S_{n+3} + \frac{\Delta\beta_n\dfrac{\Delta S_{n+2}}{\Delta S_n} + \Delta\left(\beta_n\dfrac{\Delta S_{n+1}}{\Delta S_n}\right)}{\Delta\left(\dfrac{\Delta\beta_n}{\Delta S_n}\right)}.
\end{aligned} \tag{2.9}$$

Remark 2.1. *If $a_n = a$ for $n = 0, 1, \ldots$, then* (2.1) *is the kernel of the Aitken's Δ^2 process* (1.1)*. Moreover, the system* (2.4) *has more than one solution, which are given by*

$$\begin{pmatrix} S \\ \lambda S \\ \lambda \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} S_{n+1} \\ S_n\Delta S_{n+1}/\Delta S_n \\ \Delta S_{n+1}/\Delta S_n \end{pmatrix}, \text{ for } \alpha \in \mathbb{R}. \tag{2.10}$$

The first element of this vector is ${}^1\mathrm{T}_n$. In addition, taking $\alpha = -\dfrac{\Delta S_n \Delta S_{n+1}}{\Delta^2 S_n}$, gives the process of Aitken as the second expression (1.3)*.*

The second sequence transformation is directly given by the system (2.2). Indeed, assuming $\beta_n \neq \lambda$ gives S, and taking $T_n = S$ we get the transformation

$$T_n = \frac{a_{n+1}S_n\lambda - a_nS_{n+1}}{a_{n+1}\lambda - a_n}, \tag{2.11}$$

that can be rewritten, as previously discussed, in a more stable form

$$\begin{aligned}
T_n &= S_n - \frac{a_n\Delta S_n}{a_{n+1}\lambda - a_n} = S_n - \frac{\beta_n\Delta S_n}{\lambda - \beta_n} \\
&= S_{n+1} - \frac{a_{n+1}\Delta S_n\lambda}{a_{n+1}\lambda - a_n} = S_{n+1} - \frac{\Delta S_n\lambda}{\lambda - \beta_n}.
\end{aligned} \tag{2.12}$$

However, we need to compute the unknown λ. Hence, we propose two approaches that give λ as the solution of a linear system. The first method is given by the solution of the system (2.4). Then, using the computed λ in (2.12), we obtain a transformation which we denote ${}^2\mathrm{T}_n$. The value of λ obtained as the solution of system (2.4) can be expressed as the following ratio of determinants

$$\begin{aligned}\lambda &= \frac{\begin{vmatrix} a_n & -a_{n+1} & a_n S_{n+1} \\ a_{n+1} & -a_{n+2} & a_{n+1} S_{n+2} \\ a_{n+2} & -a_{n+3} & a_{n+2} S_{n+3} \end{vmatrix}}{\begin{vmatrix} a_n & -a_{n+1} & a_{n+1} S_n \\ a_{n+1} & -a_{n+2} & a_{n+2} S_{n+1} \\ a_{n+2} & -a_{n+3} & a_{n+3} S_{n+2} \end{vmatrix}} \\ &= \frac{\begin{vmatrix} a_{n+1} & a_{n+2} & a_{n+3} \\ a_n & a_{n+1} & a_{n+2} \\ a_n S_{n+1} & a_{n+1} S_{n+2} & a_{n+2} S_{n+3} \end{vmatrix}}{\begin{vmatrix} a_{n+1} & a_{n+2} & a_{n+3} \\ a_n & a_{n+1} & a_{n+2} \\ a_{n+1} S_n & a_{n+2} S_{n+1} & a_{n+3} S_{n+2} \end{vmatrix}}.\end{aligned} \tag{2.13}$$

Similarly to what is done for (2.5), if we assume $a_n \neq 0$ for $n \in \mathbb{N}_0$ and $\Delta\beta_n \neq 0$, we get

$$\lambda = \frac{\Delta\left(\dfrac{\Delta\left(\beta_n S_{n+1}\right)}{\Delta\beta_n}\right)}{\Delta\left(\dfrac{\Delta S_n}{\Delta\beta_n}\right)}. \tag{2.14}$$

Now, if β_n converges to $\beta \in \mathbb{R}$, with $\beta \neq \lambda$, then $b_{n+1}\dfrac{\Delta S_{n+1}}{\Delta S_n}$ converges to λ. Thus, (2.14) can be expressed as follows

$$\lambda = b_{n+2}\frac{\Delta S_{n+2}}{\Delta S_{n+1}} + \frac{\Delta S_n \Delta\left(\beta_n\beta_{n+1}\dfrac{\Delta S_{n+1}}{\Delta S_n}\right)}{\beta_{n+1}\Delta\beta_n\Delta\left(\dfrac{\Delta S_n}{\Delta\beta_n}\right)}.$$

Hence, we can define ${}^2\mathrm{T}_n$ as

$$\begin{aligned} {}^2\mathrm{T}_n &= S_n - \frac{\beta_n \Delta S_n \Delta\left(\dfrac{\Delta S_n}{\Delta \beta_n}\right)}{\Delta\left(\dfrac{\Delta\left(\beta_n S_{n+1}\right)}{\Delta \beta_n}\right) - \beta_n \Delta\left(\dfrac{\Delta S_n}{\Delta \beta_n}\right)} \\ &= S_{n+1} - \frac{\Delta S_n \Delta\left(\dfrac{\Delta\left(\beta_n S_{n+1}\right)}{\Delta \beta_n}\right)}{\Delta\left(\dfrac{\Delta\left(\beta_n S_{n+1}\right)}{\Delta \beta_n}\right) - \beta_n \Delta\left(\dfrac{\Delta S_n}{\Delta \beta_n}\right)}. \end{aligned} \tag{2.15}$$

We can compute λ in a different way. Indeed, if we apply the forward difference operator Δ to the system (2.2) we obtain

$$\begin{aligned} \Delta S_n &= \lambda^n (a_{n+1}\lambda - a_n), \\ \Delta S_{n+1} &= \lambda^{n+1}(a_{n+2}\lambda - a_{n+1}). \end{aligned}$$

Since the unknowns λ^n, λ^{n+1} can be eliminated by division, we get the following quadratic equation for the unknown λ

$$a_{n+2}\Delta S_n \lambda^2 - a_{n+1}(\Delta S_n + \Delta S_{n+1})\lambda + a_n \Delta S_{n+1} = 0. \tag{2.16}$$

Clearly, the equation has two solutions for λ, but we cannot choose one of them a priori. Indeed, the criterion according to which we accept one of them and reject the other one, is based on λ itself. Hence, we obtain λ solving the following system obtained by (2.16)

$$\begin{aligned} a_{n+1}(\Delta S_n + \Delta S_{n+1})\lambda - a_{n+2}\Delta S_n \lambda^2 &= a_n \Delta S_{n+1}, \\ a_{n+2}(\Delta S_{n+1} + \Delta S_{n+2})\lambda - a_{n+3}\Delta S_{n+1}\lambda^2 &= a_{n+1}\Delta S_{n+2}. \end{aligned} \tag{2.17}$$

Now, setting λ and λ^2 as two unrelated unknowns, the system is linear. Hence, we can get λ as the solution of the above system. Finally, using this value of λ in (2.12) we define the transformation ${}^3\mathrm{T}_n$.

We notice that it is possible to express λ explicitly in the following way

$$\lambda = \frac{\Delta\left(\beta_n \beta_{n+1} \dfrac{\Delta S_{n+1}}{\Delta S_n}\right)}{\Delta\left(\beta_{n+1} \dfrac{\Delta S_n + \Delta S_{n+1}}{\Delta S_n}\right)}.$$

Therefore, transformation ${}^3\mathrm{T}_n$ can be equivalently stated in the following forms

$$\begin{aligned} {}^3\mathrm{T}_n &= S_n + \beta_n \frac{\Delta S_n}{\Delta S_{n+2}} \frac{\Delta\left(\beta_{n+1}\dfrac{\Delta S_n + \Delta S_{n+1}}{\Delta S_n}\right)}{\beta_n \dfrac{\Delta\beta_{n+1}}{\Delta S_{n+2}} - \beta_{n+2}\dfrac{\Delta\beta_n}{\Delta S_{n+1}}} \\ &= S_{n+1} + \frac{\Delta S_n}{\Delta S_{n+2}} \frac{\Delta\left(\beta_n\beta_{n+1}\dfrac{\Delta S_{n+1}}{\Delta S_n}\right)}{\beta_n \dfrac{\Delta\beta_{n+1}}{\Delta S_{n+2}} - \beta_{n+2}\dfrac{\Delta\beta_n}{\Delta S_{n+1}}}. \end{aligned}$$

In the numerical experiments (see Section 2.3) we will see that computing λ as the solution of the system (2.4) seems to be more accurate than obtaining it by the system (2.17).

Remark 2.2. *If $a_n = a$ for $n = 0, 1, \ldots,$ then* (2.1) *is the kernel of the Aitken's Δ^2 process* (1.1)*; see Remark 2.1. Moreover, the solutions of the system* (2.4) *are given by the equation* (2.10)*. However, the value of λ computed solving the system* (2.4) *is $\dfrac{\Delta S_{n+1}}{\Delta S_n}, \forall \alpha \in \mathbb{R}$. Thus,* (2.12) *becomes equal to* (1.3)*, i.e., from ${}^2T_n$ we recover Aitken's Δ^2 process.*

We can give the same result for transformation ${}^3\mathrm{T}_n$. Indeed, if a_n is constant, then $\lambda = \dfrac{\Delta S_{n+1}}{\Delta S_n}$ is a solution of the system (2.17)*, since the solutions of this system are*

$$\begin{pmatrix} \lambda \\ \lambda^2 \end{pmatrix} = \alpha \begin{pmatrix} \Delta S_n \\ \Delta S_n + \Delta S_{n+1} \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \text{ for } \alpha \in \mathbb{R}.$$

Then for $\alpha = \dfrac{\Delta^2 S_n}{(\Delta S_n)^2}$ we recover Aitken's process.

2.2 Convergence and Acceleration Properties

In this section we will analyze the behavior and the convergence property of the introduced transformations. The transformation with the best performances is ${}^2\mathrm{T}_n$. We will see it in Section 2.3 using numerical experiments. Thus, we decide to analyze mainly this transformation, in particular for the convergence analysis.

To begin, we recall the following well-known criterion for the convergence of some sequences. Assume that the elements of the sequence $S_n = S + a_n\lambda^n + g_n$ satisfy

$$\lim_{n\to\infty} \frac{S_{n+1} - S}{S_n - S} = \rho.$$

We call $|\rho|$ the *convergence rate* of the sequence. In particular, as introduced in Section 1.1, we say that S_n converges linearly if $0 < |\rho| < 1$, that converges logarithmic if $\rho = 1$, and that converges hyperlinearly if $\rho = 0$. If $|\rho| > 1$ then the sequence S_n diverges.

Let us consider the sequence

$$\tilde{S}_n = S + a_n\lambda^n + g_n, \quad n = 0, 1, \ldots, \tag{2.18}$$

with g_n a "noise term". This means that we want g_n such that $\tilde{S}_n$ is not "too far" from the kernel (2.1). We characterize this concept assuming that g_n is *subdominant* to $a_n\lambda^n$ as $n \to \infty$, i.e., $\lim_{n\to\infty} g_n/(a_n\lambda^n) = 0$. This also implies that the convergence rate of $\tilde{S}_n$ depends only on the term $a_n\lambda^n$.

Let us define

$$\beta = \lim_{n\to\infty} \beta_n = \lim_{n\to\infty} \frac{a_n}{a_{n+1}},$$

then, if $|\beta|$ exists and is finite we get

$$\lim_{n\to\infty} \frac{\tilde{S}_{n+1} - S}{\tilde{S}_n - S} = \frac{\lambda}{\beta}.$$

Therefore, if $|\lambda| < |\beta|$ $\tilde{S}_n$ is linearly convergent, if $\lambda = \beta$ it has logarithmic convergence, while if $|\lambda| > |\beta|$ the sequence diverges. Moreover, if $|\beta| = \infty$, then $\tilde{S}_n$ is hyperlinearly convergent for every value of λ. Hence, in the latter case convergence acceleration methods are not useful, unless $|\lambda|$ is sufficiently large. Therefore, we exclude the case in which $|\beta| = \infty$ from our analysis of the acceleration behavior of ${}^2\mathrm{T}_n$.

When we consider the sequence $\tilde{S}_n$ instead of the sequence S_n, the value of λ obtained by solving the system (2.4) depends on n, we denote it as λ_n. With the same procedure used to obtain expression (2.14), we get the following formula for λ_n

$$\lambda_n = \frac{\Delta\left(\dfrac{\Delta\left(\beta_n\tilde{S}_{n+1}\right)}{\Delta\beta_n}\right)}{\Delta\left(\dfrac{\Delta\tilde{S}_n}{\Delta\beta_n}\right)}. \tag{2.19}$$

We will find useful the following statements about the convergence of λ_n to λ, as $n \to \infty$. First, we need the following technical lemmas.

Lemma 2.3 ([14]). *If $g_n/(a_n\lambda^n) \to 0$ and β_n is bounded, then $\Delta g_n/(a_{n+1}\lambda^n) \to 0$.*

Proof.

$$\begin{aligned}\lim_{n\to\infty} \frac{\Delta g_n}{a_{n+1}\lambda^n} &= \lim_{n\to\infty} \left(\frac{g_{n+1}}{a_{n+1}\lambda^n} - \frac{g_n}{a_{n+1}\lambda^n} \right) \\ &= \lim_{n\to\infty} \left(\lambda \frac{g_{n+1}}{a_{n+1}\lambda^{n+1}} - \frac{g_n}{a_n\lambda^n}\beta_n \right) = \lambda 0 - 0 = 0.\end{aligned}$$

□

Lemma 2.4 ([14]). *Consider the sequence $\tilde{S}_n = S + a_n\lambda^n + g_n$, $n = 0, 1, \ldots$ Assume that*

1. $\lim_{n\to\infty} \dfrac{g_n}{a_n\lambda^n} = 0\,,$
2. *there exists a finite number β such that $\beta \neq \lambda$ for which $\lim_{n\to\infty} \beta_n = \lim_{n\to\infty} \dfrac{a_n}{a_{n+1}} = \beta\,.$*

Then $\lim_{n\to\infty} \beta_{n+1} \dfrac{\Delta\tilde{S}_{n+1}}{\Delta\tilde{S}_n} = \lambda\,.$

Proof.

$$\begin{aligned}\lim_{n\to\infty} \beta_{n+1} \frac{\Delta\tilde{S}_{n+1}}{\Delta\tilde{S}_n} &= \lim_{n\to\infty} \frac{a_{n+1}}{a_{n+2}} \frac{a_{n+2}\lambda^{n+1}\left(\lambda - \dfrac{a_{n+1}}{a_{n+2}} + \dfrac{\Delta g_{n+1}}{a_{n+2}\lambda^{n+1}}\right)}{a_{n+1}\lambda^n\left(\lambda - \dfrac{a_n}{a_{n+1}} + \dfrac{\Delta g_n}{a_{n+1}\lambda^n}\right)} \\ &= \lim_{n\to\infty} \lambda \frac{\lambda - \beta_{n+1} + \dfrac{\Delta g_{n+1}}{a_{n+2}\lambda^{n+1}}}{\lambda - \beta_n + \dfrac{\Delta g_n}{a_{n+1}\lambda^n}} = \lambda \frac{\lambda - \beta + 0}{\lambda - \beta + 0} = \lambda.\end{aligned}$$

We remark that $\Delta g_n/(a_{n+1}\lambda^n)$ and $\Delta g_{n+1}/(a_{n+2}\lambda^{n+1})$ converge to zero by Lemma 2.3. □

We recall that the meaning of hypothesis 1 is that the sequence is "not too far" from the kernel of the transformation. Moreover, notice that if $\beta = \lambda$ (hypothesis 2 is not satisfied), then $\tilde{S}_n$ is logarithmically convergent. The case in which β does not exist is not easily interpreted. In Section 2.3.5 we discuss several numerical examples related to these situations.

Let us define the sequence

$$\gamma_n = \frac{a_{n+2}^2 - a_{n+1}a_{n+3}}{a_{n+1}^2 - a_n a_{n+2}}.$$

Then we can give the following theorem.

Theorem 2.5 ([14])**.** *The sequence (λ_n) converges to λ if the following conditions are satisfied:*

1. $\lim_{n\to\infty} \frac{g_n}{a_n\lambda^n} = 0\,,$
2. *there exists* $\beta \in \mathbb{R}$ *such that* $\lim_{n\to\infty} \beta_n = \beta\,,$
3. *there exists* $\gamma \in \mathbb{R}$ *such that* $\lim_{n\to\infty} \gamma_n = \gamma\,,$
4. *λ, β and γ are such that $\beta \neq \lambda$ and $\lambda - \beta^3\gamma \neq 0$.*

Proof. From (2.19) we get the following formulas for λ_n

$$\begin{aligned}\lambda_n &= \frac{\beta_{n+2}\,\Delta\beta_n\,\Delta\tilde{S}_{n+2} - \beta_n\,\Delta\beta_{n+1}\,\Delta\tilde{S}_{n+1}}{\Delta\beta_n\,\Delta\tilde{S}_{n+1} - \Delta\beta_{n+1}\,\Delta\tilde{S}_n} \\ &= \frac{\beta_{n+2}\dfrac{\Delta\tilde{S}_{n+2}}{\Delta\tilde{S}_{n+1}} - \beta_n\dfrac{\Delta\beta_{n+1}}{\Delta\beta_n}}{1 - \dfrac{\Delta\beta_{n+1}}{\Delta\beta_n}\dfrac{\Delta\tilde{S}_n}{\Delta\tilde{S}_{n+1}}}\,.\end{aligned}$$

Moreover, we rewrite the term $\Delta\beta_{n+1}/\Delta\beta_n$ as

$$\frac{\Delta\beta_{n+1}}{\Delta\beta_n} = \frac{a_{n+1}}{a_{n+3}}\left(\frac{a_{n+2}^2 - a_{n+1}a_{n+3}}{a_{n+1}^2 - a_n a_{n+2}}\right) = \beta_{n+1}\beta_{n+2}\gamma_n.$$

Now, using Lemmas 2.3 and 2.4 gives

$$\begin{aligned}\lim_{n\to\infty}\lambda_n &= \lim_{n\to\infty}\frac{\beta_{n+2}\dfrac{\Delta\tilde{S}_{n+2}}{\Delta\tilde{S}_{n+1}}-\beta_n\beta_{n+1}\beta_{n+2}\gamma_n}{1-\beta_{n+1}\beta_{n+2}\gamma_n\dfrac{\Delta\tilde{S}_n}{\Delta\tilde{S}_{n+1}}}\\ &= \lim_{n\to\infty}\frac{\beta_{n+2}\dfrac{\Delta\tilde{S}_{n+2}}{\Delta\tilde{S}_{n+1}}-\beta_n\beta_{n+1}\beta_{n+2}\gamma_n}{1-\beta_{n+1}^2\beta_{n+2}\gamma_n\left(\beta_{n+1}\dfrac{\Delta\tilde{S}_{n+1}}{\Delta\tilde{S}_n}\right)^{-1}}\\ &= \frac{\lambda-\beta^3\gamma}{1-\dfrac{\beta^3\gamma}{\lambda}}=\lambda\frac{\lambda-\beta^3\gamma}{\lambda-\beta^3\gamma}=\lambda.\end{aligned}$$

□

Note that β_n and γ_n only depend on the sequence (a_n). Then, since in our study (a_n) is considered to be known, we can check if sequences (β_n) and (γ_n) have a limit, and, if we are able to compute them, we can know the values of λ for which the estimate λ_n may not converge to the correct limit. Moreover, we remark that γ_n may not be well defined if $a_{n+1}^2 - a_n a_{n+2} = 0$ for some n. However, if $a_{n+2}^2 - a_{n+1}a_{n+3} \neq 0$, we can skip this iteration, and compute the following one. If both the numerator and the denominator of γ_n are equal to zero, then also the denominator of λ as expressed in (2.13) is equal to zero. Hence, if γ_n is not well defined, λ_n cannot be computed.

Now, we can state some results on the acceleration properties of ${}^2\mathrm{T}_n$. Let us consider first the **convergent case**.

If $\tilde{S}_n \to S$, then we get

$$\begin{aligned}{}^2\mathrm{T}_n &= \frac{(a_n - a_{n+1}\lambda_n)\tilde{S}_{n+1} + a_{n+1}(\tilde{S}_{n+1}-\tilde{S}_n)\lambda_n}{a_n - a_{n+1}\lambda_n}\\ &= \frac{a_n\tilde{S}_{n+1} - a_{n+1}\tilde{S}_n\lambda_n}{a_n - a_{n+1}\lambda_n}\\ &= \frac{\frac{a_n}{a_{n+1}}\tilde{S}_{n+1} - \tilde{S}_n\lambda_n}{\frac{a_n}{a_{n+1}} - \lambda_n}.\end{aligned}$$

Hence,

$${}^2\mathrm{T}_n - S = \frac{\beta_n(\tilde{S}_{n+1}-S) - (\tilde{S}_n - S)\lambda_n}{\beta_n - \lambda_n}. \tag{2.20}$$

Using (2.20) we immediately deduce the following theorem.

Theorem 2.6 ([14]). *Transformation* ${}^2\mathrm{T}_n$ *converges to* S *under the following conditions:*

1. $\lim_{n\to\infty} \tilde{S}_n = S$,
2. *there exist* $N \in \mathbb{N}$ *and* $\delta > 0$ *such that* $|\lambda_n - \beta_n| > \delta$ *for every* $n > N$.

It remains to prove that transformation ${}^2\mathrm{T}_n$ accelerates the convergence of sequences of the form of (2.18). First, we need the following lemma.

Lemma 2.7 ([14]). *If* $g_n/(a_n\lambda^n) \to 0$, *then*

$$\lim_{n\to\infty} \beta_n \frac{\tilde{S}_{n+1} - S}{\tilde{S}_n - S} = \lambda\,.$$

Proof.

$$\begin{aligned}
\lim_{n\to\infty} \beta_n \frac{\tilde{S}_{n+1} - S}{\tilde{S}_n - S} &= \lim_{n\to\infty} \beta_n \frac{a_{n+1}\lambda^{n+1} + g_{n+1}}{a_n\lambda^n + g_n} \\
&= \lim_{n\to\infty} \beta_n \frac{a_{n+1}}{a_n}\lambda \frac{1 + \dfrac{g_{n+1}}{a_{n+1}\lambda^{n+1}}}{1 + \dfrac{g_n}{a_n\lambda^n}} \\
&= \lim_{n\to\infty} \lambda \frac{1 + \dfrac{g_{n+1}}{a_{n+1}\lambda^{n+1}}}{1 + \dfrac{g_n}{a_n\lambda^n}} = \lambda.
\end{aligned}$$

□

Theorem 2.8 ([14]). *Under the assumptions of Theorem 2.5, transformation* ${}^2\mathrm{T}_n$ *accelerates the convergence of the sequence (2.18).*

Proof. By (2.20) we get

$$\frac{{}^2\mathrm{T}_n - S}{\tilde{S}_n - S} = \frac{\beta_n(\tilde{S}_{n+1} - S) - (\tilde{S}_n - S)\lambda_n}{(\beta_n - \lambda_n)(\tilde{S}_n - S)} = \frac{\beta_n \frac{\tilde{S}_{n+1}-S}{\tilde{S}_n - S} - \lambda_n}{\beta_n - \lambda_n}.$$

Moreover, Theorem 2.5 gives $\lambda_n \to \lambda$. Therefore, assuming $\beta \neq \lambda$ implies

$$\lim_{n\to\infty} \frac{{}^2\mathrm{T}_n - S}{\tilde{S}_n - S} = 0.$$

□

Notice that the theorem holds for any estimate $\tilde{\lambda}_n$ converging to λ. In addition, the convergence of λ_n to λ is the key to prove acceleration and convergence.

Finally, let us consider a **divergent sequence** $\tilde{S}_n$

By equation (2.20) we get

$$\begin{aligned} {}^2\mathrm{T}_n - S &= \frac{\beta_n(a_{n+1}\lambda^{n+1} + g_{n+1}) - \lambda_n(a_n\lambda^n + g_n)}{\beta_n - \lambda_n} \\ &= \frac{a_n\lambda^n(\lambda - \lambda_n) + \beta_n g_{n+1} - \lambda_n g_n}{\beta_n - \lambda_n} \\ &= \frac{a_n\lambda^n(\lambda - \lambda_n)}{\beta_n - \lambda_n} + \frac{\beta_n g_{n+1} - \lambda_n g_n}{\beta_n - \lambda_n}. \end{aligned}$$

It is easy to give assumptions under which the term

$$\frac{\beta_n g_{n+1} - \lambda_n g_n}{\beta_n - \lambda_n}$$

of the last equation converges to zero. Nevertheless, since the sequence $a_n\lambda^n$ is divergent there are not meaningful conditions for which the term

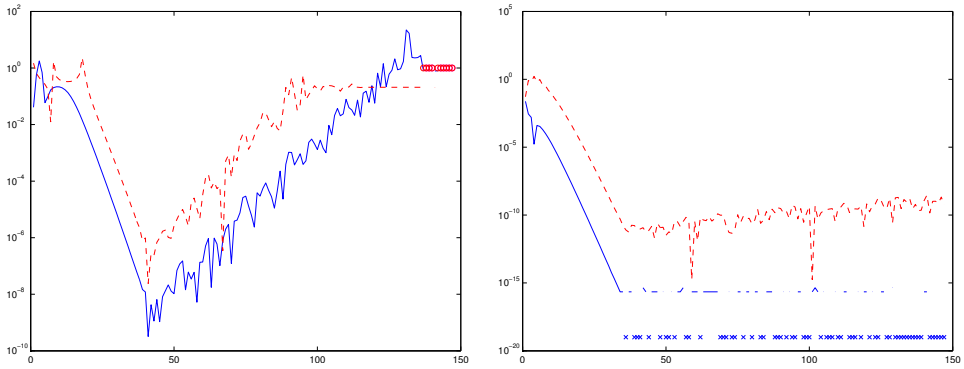
$$\frac{a_n\lambda^n(\lambda - \lambda_n)}{\beta_n - \lambda_n}$$

does not diverge. In particular, hypotheses of Theorem 2.5 gives $\lambda - \lambda_n \to 0$, but they do not ensures the convergence of ${}^2\mathrm{T}_n$. In general, our numerical experiments show that transformation ${}^2\mathrm{T}_n$ does not converge when the sequence diverges.

However, we can give some results about semi-convergence in the following remark. We have semi-convergence when a sequence has a convergent behavior at the first iterations, but then it diverges. To our knowledge the concept of semi-convergence was introduced by Stieltjes [29]. For a review about semi-convergence we refer to [36, Appendix E].

Remark 2.9. *When λ_n rapidly converges to λ and $a_n\lambda^n$ does not diverge too quickly at the beginning, ${}^2\mathrm{T}_n$ may have a semi-convergent behavior, we will see this in the numerical experiments in Section 2.3.*

We last remark that if $\lambda_m = \lambda$ for a certain m, then $T_m = S + \varepsilon$, where $\varepsilon = (\beta_m g_{m+1} - \lambda_m g_m)/(\beta_m - \lambda_m)$ can be supposed to be very small. This may explain why in the numerical experiments we sometimes have values of ${}^2\mathrm{T}_n$ which are very close to S, even if the transformation generally diverges.



(a) sequence (2.21) (b) sequence (2.22)

Figure 2.1: Comparison of the absolute value of the error in the estimate of λ solving systems (2.4) (solid), and (2.17) (dashed).

2.3 Numerical Experiments

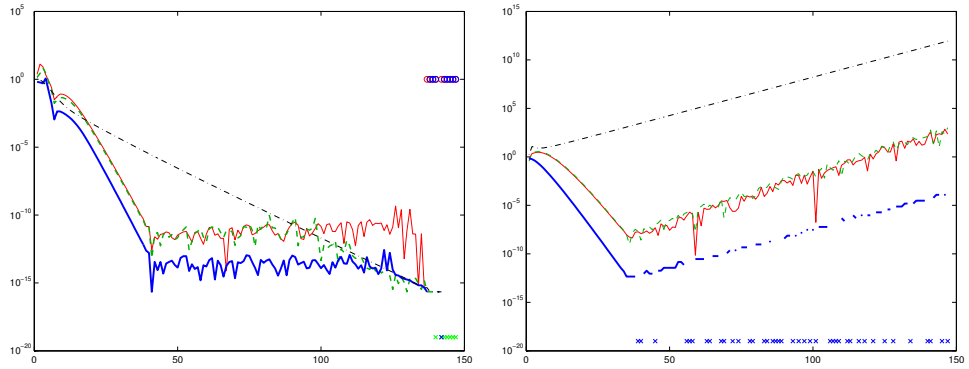
In this Section, we will present the numerical experiments reported in [14]. We first discuss the approximation of λ and the results of the transformations ${}^1\mathrm{T}_n$, ${}^2\,\mathrm{T}_n$ and ${}^3\mathrm{T}_n$. Secondly, we compare the best of the three transformations, ${}^2\mathrm{T}_n$, with other well-known and classical transformations and with the transformations presented in [13]. Finally, we will consider some cases in which the convergence of λ_n to λ is not ensured and so ${}^2\mathrm{T}_n$ could fail, see Section 2.2.

The experiments were obtained using `Matlab 7.12.0`. While computing λ or ${}^2\mathrm{T}_n$ by solving a linear system, sometimes a singular matrix appears. Then, we mark this with a circle $\circ$ in the plot at the corresponding iteration. Moreover, we mark with a $\times$ the iterations in which ${}^2\mathrm{T}_n$ or λ are computed at machine precision. We remark that whenever we compute λ as the solution of systems (2.4) or (2.17), we use the `Matlab` backslash command \.

2.3.1 Estimation of λ

Let us consider two sequences. The first one is linearly convergent and satisfies the condition of Theorem 2.8

$$S_n = 1 + \log\left(1 + \frac{1}{n}\right)\left(\frac{4}{5}\right)^n + e^{-n}(1 + n^2). \tag{2.21}$$



(a) sequence (2.21) (b) sequence (2.22)

Figure 2.2: Comparison of $|S - S_n|$ (dash-dotted) and $|S - T_n|$ for transformations ${}^1\mathrm{T}_n$ (dashed), ${}^2\mathrm{T}_n$ (bold) and ${}^3\mathrm{T}_n$ (solid).

It has convergence rate $\rho = \frac{4}{5}$ and $\beta = \lim_{n\to\infty} a_n/a_{n+1} = 1$. The second one is divergent and such that ${}^2T_n$ is expected to semi-converge, accordingly to Remark 2.9. Indeed, it is alternating, divergent and has $\beta = \lim_{n\to\infty} a_n/a_{n+1} = 1$.

$$S_n = 1 + \left[10\sin\left(\pi\left(1+\frac{1}{n^2}\right)\right) + \right. \\ \left. +2\cos\left(\pi\left(1+\frac{1}{n^2}\right)\right)\right]\left(-\frac{6}{5}\right)^n + e^{-n}(1+n^2). \tag{2.22}$$

In Figure 2.1, the solid line is the absolute error of the estimate of λ obtained by solving the system (2.4), while the dashed line is the corresponding absolute error obtained by solving (2.17). The first system seems to give a better approximation than the second one. In particular, in certain cases (see Figure 2.1b), the solution of the system (2.4) rapidly reaches machine precision. Nevertheless, for converging sequences we have $\Delta S_n \to 0$. Hence, rounding errors appear in the solution of the systems (2.4) and (2.17). Indeed, as we can see in Figure 2.1a, both approximations reach a good precision before diverging.

2.3.2 Comparison between the proposed transformations

When S_n is convergent, we consider the best transformation the one that converges to S with fewest iterations and good precision. However,

for diverging sequences, we expect a good transformation to perform a semi-convergence behavior as shown in Remark 2.9; see in particular [13].

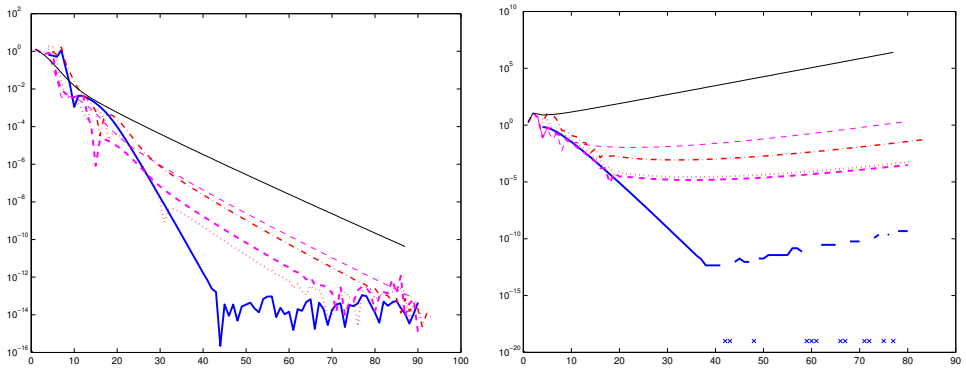
We present the comparison between the performance of the three transformations in Figure 2.2. The results are obtained using for ${}^1\mathrm{T}_n$ the last expression of (2.7) with denominator $D_n = a_{n+3}S_{n+2}(a_{n+1}^2 - a_n a_{n+2}) + a_{n+2}S_{n+1}(a_n a_{n+3} - a_{n+1}a_{n+2}) + a_{n+1}S_n(a_{n+2}^2 - a_{n+1}a_{n+3})$. While, for both ${}^2\mathrm{T}_n$, ${}^3\mathrm{T}_n$ we use the third formula of (2.12), with λ the solution of the system (2.4) and (2.17) respectively. The absolute errors $|S - {}^i\mathrm{T}_n|$, for $i = 1, 2, 3$, are plotted in dashed, bold, solid lines, respectively, while the dash-dotted line correspond to $|S - S_n|$. The three transformations accelerate the convergence of the sequence. Moreover, their performance is good even when the estimate of λ is not, as we can see looking at figures 2.1a and 2.2a. Clearly, the best result is the one given by transformation ${}^2\mathrm{T}_n$. Note that for every n all the transformations use the same sequence terms $S_n, S_{n+1}, S_{n+2}, S_{n+3}$, hence their computational cost is almost the same. However, assuming that the sequences are known, the time needed for computing the first 100 values in the plots of figure 2.2a are 0.0025 seconds for ${}^1\mathrm{T}_n$, 0.0076 seconds for ${}^2\mathrm{T}_n$ and 0.0072 seconds for ${}^3\mathrm{T}_n$. While the time needed for the first 100 values in the plots of figure 2.2b are 0.0032 seconds for ${}^1\mathrm{T}_n$, 0.0083 seconds for ${}^2\mathrm{T}_n$ and 0.0070 for ${}^3\mathrm{T}_n$. Clearly computing the values of ${}^1\mathrm{T}_n$ is faster than the other transformations since it has not to solve a linear system.

2.3.3 Comparison with other transformations

We compare transformation ${}^2\mathrm{T}_n$ with other well-known transformations, and with the transformations introduced by Brezinski and Redivo-Zaglia in [13].

In Figure 2.3 we plot the absolute error, $|S - T_n|$, for every iteration n, with T_n one of the following transformations:

- transformation ${}^2\mathrm{T}_n$, plotted in solid bold line, which uses four terms of the sequence;
- ε-algorithm ($\varepsilon_{2k}^{(n)}$) with $k = 2$ (see Section 1.2), plotted in dashed bold line, which uses five terms;
- Aitken's Δ^2 process uses three terms and it is plotted in dashed line. Since it is known that $T_n = \varepsilon_2^{(n)}$ for $n = 1, 2, \ldots$ (see Section 1.2), where T_n is the Aitken's Δ^2 process, we use the ε-algorithm for the computation of the Aitken's Δ^2 process;

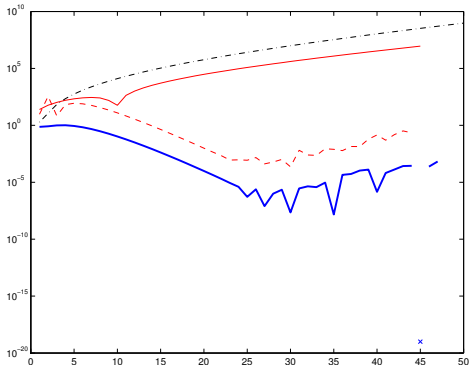


(a) sequence (2.21) (b) sequence (2.22)

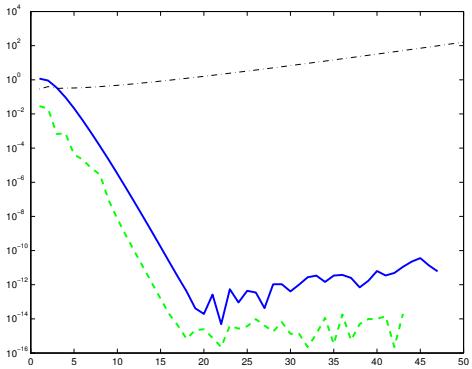
Figure 2.3: Comparison of $|S - S_n|$ (solid) and $|S - T_n|$ values using Aitken's Δ^2 process (dashed), $\varepsilon_4^{(n)}$ (dashed bold), $\theta_2^{(n)}$ (dash-dotted), Levin type transformation (dotted) and transformation ${}^2\mathrm{T}_n$ (bold).

- algorithm $\theta_2^{(n)}$ (see Section 1.3), which uses four terms and is plotted in dash-dotted line;
- Levin type transformation $\mathcal{L}_k^{(n)}(\beta, S_n, \omega_n)$ (see Section 1.4), plotted with dots. For ω_n we use the formula that gives the u-transformation and we set $k = 3$ so that the transformation uses four terms. The parameter b is chosen equal to 1, that is the optimal choice for our sequences following the procedure described in [1]. However, other values of b give similar results in our experiments.

Analyzing Figure 2.3a we first notice that ${}^2\mathrm{T}_n$ converges faster than the other transformations. Moreover, in Figure 2.3b we consider the divergent sequence (2.22). As we can see, all the transformations semi-converge, however ${}^2\mathrm{T}_n$ is the one who reaches the highest accuracy, before diverging. We underline that we expected this performance by ${}^2\mathrm{T}_n$ since the transformation was built from the kernel $S_n = S + a_n\lambda^n$, hence it should have a good performance for sequences of the type of (2.18), as the one we are considering. Assuming that the sequences are known, the time needed for computing the values in the plots of figure 2.3a are 0.0082 seconds for ${}^2\mathrm{T}_n$, 0.0092 seconds for Aitken's process, 0.0146 seconds for the ε-algorithm, 0.0025 seconds for the θ-algorithm, and 0.0037 seconds for the Levin type transformation. While the time needed for the values in the plots of figure 2.3b are 0.0081 seconds for ${}^2\mathrm{T}_n$, 0.0070 seconds for Aitken's process, 0.0120 seconds for the ε-algorithm, 0.0026



(a) Comparison of $|S-S_n|$ (dash-dotted) and $|S-T_n|$ using the transformations (2.25) (dashed), (2.26) (solid) and transformation ${}^2\mathrm{T}_n$ (bold) on sequence (2.28).



(b) Comparison of $|S-S_n|$ (dash-dotted) and $|S-T_n|$ using transformations (2.27) (bold-dashed) and transformation ${}^2\mathrm{T}_n$ (bold) on sequence (2.29).

Figure 2.4

seconds for the θ-algorithm, and 0.0044 seconds for the Levin type transformation.

We now compare ${}^2\mathrm{T}_n$ with the transformations proposed by Brezinski and Redivo-Zaglia in [13]. They introduce two kernels consisting of sequences of the form

$$S_n = S + (a + bx_n)\lambda^n, \quad n = 0, 1, \ldots, \tag{2.23}$$

or

$$S_n = S + (a + bx_n)^{-1}\lambda^n, \quad n = 0, 1, \ldots, \tag{2.24}$$

where S, a, b and λ are unknown parameters and (x_n) a known sequence. Then they define these following two transformations for sequences in the first kernel (2.23)

$${}^4T_n = S_{n+1} - \frac{\Delta S_{n+1} - \lambda^2 r_n \Delta S_n}{(\lambda r_n - 1)(1 - \lambda)}, \text{ with } r_n = \frac{\Delta x_{n+1}}{\Delta x_n}, \tag{2.25}$$

and

$${}^5T_n = S_{n+1} + \frac{\Delta S_{n+1} - \lambda^2 \Delta S_n}{(1 - \lambda)^2}. \tag{2.26}$$

In both cases λ is obtained by solving a linear systems.

The sequences in the form (2.24) are treated using the following transformation

$${}^6T_n = \frac{N_n}{D_n}, \tag{2.27}$$

with

$$N_n = \lambda^2 S_{n+1}(S_{n+2} - S_n) - 2\lambda(S_{n+3}S_{n+1} - S_{n+2}S_n) + S_{n+2}(S_{n+3} - S_{n+1}),$$
$$D_n = \lambda^2(S_{n+2} - S_n) - 2\lambda(S_{n+3} - S_{n+2} + S_{n+1} - S_n) + (S_{n+3} - S_{n+1}).$$

The unknowns λ and λ^2 are computed by solving the following linear system with unknowns $\lambda, \lambda^2, \lambda^2 S, \lambda S, S$

$$\begin{aligned} &\lambda^2 S_{n+1+i}(S_{n+2+i} - S_{n+i}) - 2\lambda(S_{n+1+i}S_{n+3+i} - S_{n+i}S_{n+2+i}) \\ &-\lambda^2 S(S_{n+2+i} - S_{n+i}) + 2\lambda S(S_{n+3+i} - S_{n+2+i} + S_{n+1+i} - S_{n+i}) \\ &-S(S_{n+3+i} - S_{n+1+i}) = -S_{n+2+i}(S_{n+3+i} - S_{n+1+i}), \quad i = 0, \ldots, 4. \end{aligned}$$

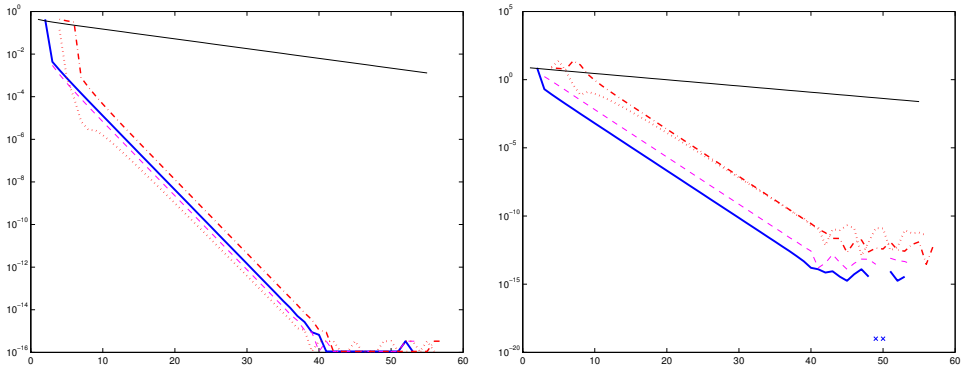
We compare transformation ${}^2\mathrm{T}_n$ with transformations ${}^4\mathrm{T}_n$,${}^5\,\mathrm{T}_n$ and ${}^6\mathrm{T}_n$ on the same sequences used in [13], that are

$$S_n = S + \lambda^n(2 - n^{\frac{7}{2}}) + e^{-n}(1 + n^2), \text{ with } \lambda = \frac{23}{20}\,, \tag{2.28}$$

and

$$S_n = S + \lambda^n \frac{1}{\left(2 + \frac{11}{10}n\right)} + \left(\frac{1}{10}\right)^n n^{\frac{5}{2}}, \text{ with } \lambda = -\frac{6}{5}\,. \tag{2.29}$$

In Figure 2.4a the dashed line is the absolute error of transformation (2.25) (which uses 6 terms), the solid line is the absolute error of transformation (2.26) (which uses 5 terms) and the bold line the absolute error of transformation ${}^2\mathrm{T}_n$ (which uses 4 terms). In Figure 2.4b the bold-dashed line is the absolute error of transformation (2.27) (which uses 8 terms) and the bold line is the absolute error of transformation ${}^2\mathrm{T}_n$. Discussing the results it is important to remark that in our computations we took as known the sequence $a_n = a + bx_n$ or $a_n = (a + bx_n)^{-1}$, whereas, in [13] a and b are unknowns and only x_n is known. This may explain why the bold line of ${}^2\mathrm{T}_n$ seems to converge faster than the other transformations. Nevertheless, even if this holds for transformations (2.25) and (2.26), in Figure 2.4b the transformation (2.27) produces better results than ${}^2\mathrm{T}_n$. Finally, assuming that the sequences are known, the time needed for computing the values in the plots of figure 2.4a are 0.0042 seconds for ${}^2\mathrm{T}_n$, 0.0033 seconds for ${}^4\mathrm{T}_n$ and 0.0020 seconds for ${}^5\mathrm{T}_n$. While the time needed for the values in the plots of figure 2.4b are 0.0023 seconds for ${}^2\mathrm{T}_n$ and 0.0019 seconds for ${}^6\mathrm{T}_n$.



(a) z = 0.9 (b) z = -0.9

Figure 2.5: Digamma function. Comparison of $|\mathcal{Z}(z) - \mathcal{Z}_n(z)|$ (solid) and $|S - T_n|$ values using Aitken's Δ^2 process (dashed), $\theta_2^{(n)}$ (dash-dotted), Levin type transformation (dotted) and transformation ${}^2\mathrm{T}_n$ (bold).

2.3.4 Computation of the digamma function

In this subsection we show some results on the acceleration through transformation ${}^2\mathrm{T}_n$ of a sequence approximating the digamma function. We consider the following power series representation of the psi or digamma function (see [2])

$$\psi(1+z) = -\gamma + z\,\mathcal{Z}(z) \tag{2.30}$$

$$\mathcal{Z}(z) = \sum_{\nu=0}^{\infty} \zeta(\nu+2)(-z)^{\nu} \tag{2.31}$$

where γ is the Euler's constant and $\zeta(\nu+2)$ is the Riemann zeta function; we refer to, e.g., [2, Equations (6.1.3) and (23.2.1)] respectively or [34, Equation 1.2]. Notice that the series in (2.31) converges for $|z| < 1$. Following the path described in [34], we rewrite (2.31) as

$$\mathcal{Z}(z) = \mathcal{Z}_n(z) + \mathcal{R}_n(z), \tag{2.32}$$

$$\mathcal{Z}_n(z) = \sum_{\nu=0}^{n} \zeta(\nu+2)(-z)^{\nu}, \tag{2.33}$$

$$\mathcal{R}_n(z) = (-z)^{n+1} \sum_{\nu=0}^{\infty} \zeta(n+\nu+3)(-z)^{\nu}; \tag{2.34}$$

see [34, Equation 2.1]. We can increase the convergence rate of the sequence $\mathcal{Z}(z)$ transforming the truncation errors $\mathcal{R}_n(z)$ into other truncation errors $\mathcal{R}'_n(z)$ with better numerical properties. Replacing the zeta

functions $\zeta(n+\nu+3)$ in (2.34) by their Dirichlet series and interchange the order of summations,

$$\mathcal{Z}_n(z) = \mathcal{Z}(z) - (-1)^{n+1} \sum_{m=0}^{\infty} \frac{[z/(m+1)]^{n+1}}{(m+1)(m+z+1)}; \tag{2.35}$$

see [34, Equation 2.2]. By the preceding equation we can see that the partial sums $\mathcal{Z}_n(z)$ are a special case of the class of sequences

$$s_n = s + (-1)^{n+1} \sum_{j=1}^{\infty} c_j (q_j)^{n+1},$$

with $q_j = z/j$ and $c_j = -1/[j(j+z)]$; see [34, Equation 2.3]. As done in [34, Equation 2.4], we assume that $q_1, q_2, \ldots$ have all the same sign and are ordered in magnitude, i.e,

$$1 > |q_1| > |q_2| > ... > |q_\ell| > |q_{\ell+1}| > ... \geq 0.$$

Whereas, the c_j are unspecified coefficients.

The digamma function expressed as in (2.35) is of the type of (2.18), with

$$\begin{aligned} \tilde{S}_n &= \mathcal{Z}_n(z), \\ S &= \mathcal{Z}(z), \\ a_n &= (-1)^n \frac{z}{z+1}, \\ \lambda &= z, \\ g_n &= (-1)^n \sum_{m=1}^{\infty} \frac{[z/(m+1)]^{n+1}}{(m+1)(m+z+1)}. \end{aligned}$$

In this numerical experiment the value of λ is known. Hence, there is no need to approximate it for transformation ${}^2\mathrm{T}_n$. Therefore, to compute ${}^2\mathrm{T}_n$ we only need two terms of the sequence. For this reason in the numerical experiments of the digamma function we will not consider the algorithm $\varepsilon_4^{(n)}$.

Figure 2.5 shows that transformation ${}^2\mathrm{T}_n$ has a similar behavior, or slightly better, than Aitken's Δ^2 process, $\theta_2^{(n)}$ algorithm and u-transformation. In particular, when $z < 0$ transformation ${}^2\mathrm{T}_n$ reaches a better precision. Assuming that the sequences are known, the time needed for computing the values in the plots of figure 2.5a are 0.0011 seconds for ${}^2\mathrm{T}_n$, 0.0024 seconds for Aitken's process, 0.0016 seconds for the θ-algorithm, and 0.0023 seconds for the Levin type transformation. We underline that since in this case we do not need to compute λ the computation of ${}^2\mathrm{T}_n$ values is faster than in the previous examples.

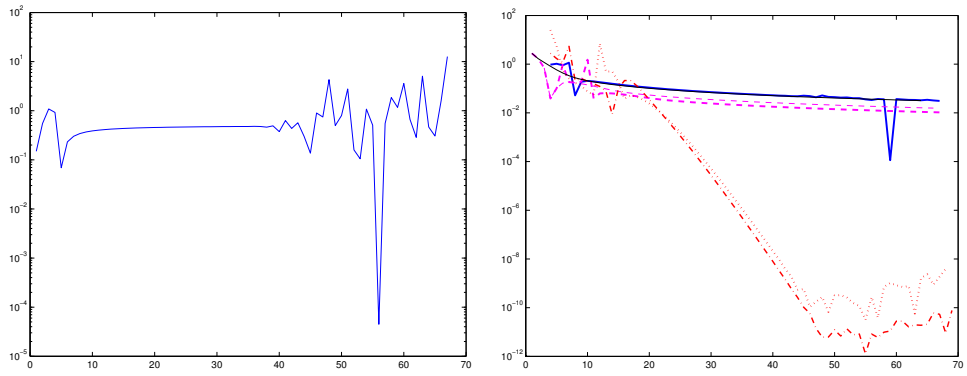


Figure 2.6: On the left, the errors in the estimate of λ (obtained by system (2.4)) for sequence (2.36). On the right, the values $|S - S_n|$ (solid) are compared with the errors obtained using Aitken's Δ^2 process (dashed), $\varepsilon_4^{(n)}$ (dashed bold), $\theta_2^{(n)}$ (dash-dotted), Levin type transformation (dotted) and transformation ${}^2\mathrm{T}_n$ (bold).

2.3.5 Problematic cases

We will show several examples of sequences for which at least one of the assumptions of Theorem 2.5 does not hold. As we have already discussed, if $\beta = \pm\infty$, then $\tilde{S}_n$ converges very rapidly. Hence, this case will not be taken into account. We consider sequences of the form

$$S_n = 1 + a_n\lambda^n + g_n,$$

with $S = 1$ and $g_n = (1 + n^2)e^{-n}$ a sequence converging to zero and subdominant to $a_n\lambda^n$.

In figures 2.6, 2.7 and 2.8 the curves on the left are the absolute error of the estimate of λ obtained by solving system (2.4). On the right, we plot the comparison between the absolute errors of respectively the transformations ${}^2\mathrm{T}_n$, Aitken's Δ^2 process, ε-algorithm, $\theta_2^{(n)}$ algorithm and u-transformation; see subsection 2.3.3.

In the first two examples we assume $\beta = \lambda$. As shown in Section 2.2, a convergent sequence of the kind of (2.18) for which $\beta = \lambda$ has a logarithmic convergence. It is well-known that Aitken's Δ^2 process and ε-algorithm are not able to accelerate logarithmically convergent sequences; see, e.g., [12]. However, we try ${}^2\mathrm{T}_n$ on a logarithmically convergent sequence to see if its behavior is similar to the one of Aitken's Δ^2 process and Wynn's ε-algorithm.

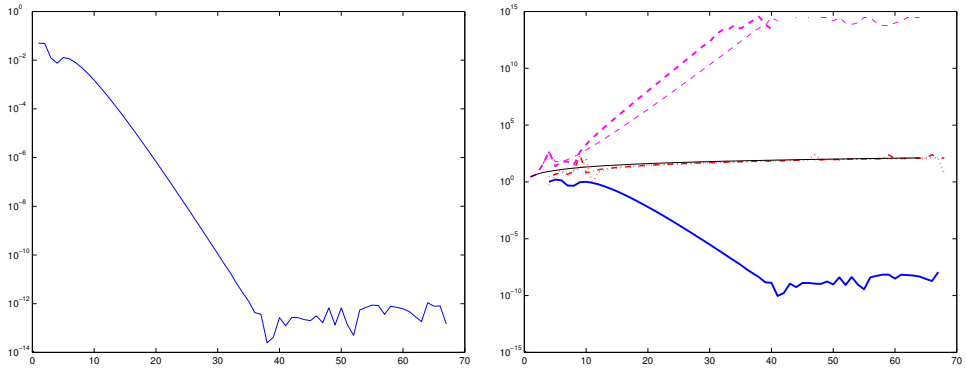


Figure 2.7: On the left, the errors in the estimate of λ (obtained by system (2.4)) for sequence (2.37). On the right, the values $|S - S_n|$ (solid) are compared with the errors obtained using Aitken's Δ^2 process (dashed), $\varepsilon_4^{(n)}$ (dashed bold), $\theta_2^{(n)}$ (dash-dotted), Levin type transformation (dotted) and transformation ${}^2\mathrm{T}_n$ (bold).

Setting $a_n = (5/4)^{n+\frac{16}{5}}/n$ and $\lambda = 4/5$, we get the sequence

$$S_n = 1 + \frac{1}{n}\left(\frac{5}{4}\right)^{n+\frac{16}{5}}\left(\frac{4}{5}\right)^n + (1+n^2)e^{-n}. \tag{2.36}$$

Notice that in this case $\beta = \lambda$ and $\gamma = \beta^{-2}$. As we expect, In Figure 2.6a λ_n does not converge to λ. Moreover, in Figure 2.6b ${}^2\mathrm{T}_n$ converges to S, but it does not accelerate the convergence, as happened for Aitken's Δ^2 process and ε-algorithm. Therefore, ${}^2\mathrm{T}_n$ seems to inherit this inability. Finally, $\theta_2^{(n)}$ and u-transformation perform a good acceleration.

Let us consider a diverging sequence $\tilde{S}_n$ for which $\lambda = \beta$. It cannot be a sequence with alternating sign. Indeed, it can be rewritten as $\tilde{S}_n = (-1)^n a_n \lambda^n + g_n$, with a_n and λ positive for every n. Then, $\beta = \lim_{n\to\infty} -a_n/a_{n+1} < 0$, while $\lambda > 0$. Hence, $\tilde{S}_n$ must be a sequence with positive terms. We underline that the summation of such sequences can be very difficult; for a further discussion about problems in the summation of this kind of sequence we refer to, e.g., [35, pp. 15-17].

Therefore, if $a_n = n\,(5/4)^{n+\frac{16}{5}}$, we get the following divergent sequence

$$S_n = 1 + n\left(\frac{5}{4}\right)^{n+\frac{16}{5}}\left(\frac{4}{5}\right)^n + (1+n^2)e^{-n}. \tag{2.37}$$

For this sequence we still have $\beta = \lambda$ and $\gamma = \beta^{-2}$. Thus, the fourth assumption of Theorem 2.5 is not satisfied. However, as we can see in

Figure 2.7 transformation ${}^2\mathrm{T}_n$ semiconverges, which agrees with Remark 2.9. In particular, λ_n converges to λ. Then, the assumption in Theorem 2.5 appears to be sufficient but not necessary. Moreover, the Aitken's process and ε-algorithm diverge, $\theta_2^{(n)}$ and u-transformation diverge at the same rate of the sequence.

Summarizing, when $\tilde{S}_n$ is a convergent sequence, $\beta = \lambda$ if and only if $\tilde{S}_n$ converges logarithmically. In this case Sequence (2.36) is an example of a logarithmically convergent sequence that ${}^2T_n$ is not able to accelerate. This result is consistent, since there is no sequence transformation that can accelerate the convergence of all logarithmically convergent sequences, see Section 1.1 and [16, 17]. Moreover, we have shown that if $\beta = \lambda$ then the sequence is definitely positive. Thus, if the sequence is divergent, then we are summing a monotone sequence, that is a class of sequences difficult to treat.

Finally, we define a different kind of sequence where

$$a_n = \frac{3}{2} + \frac{(-1)^n}{2},$$

which alternatively assumes the values 1 and 2. That is

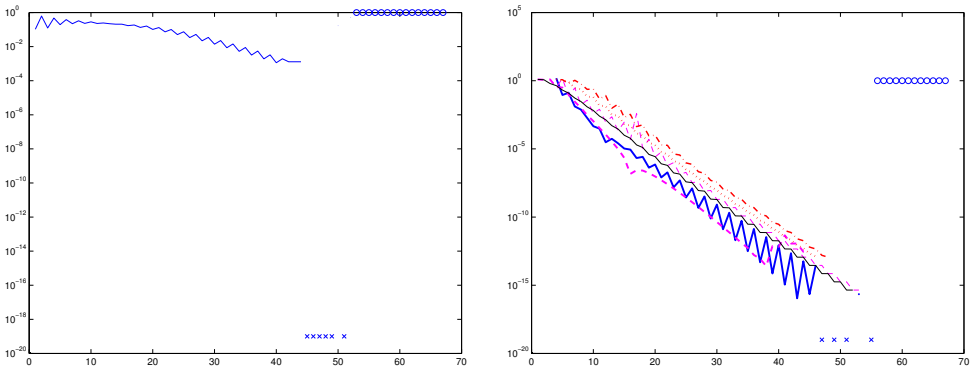
$$S_n = 1 + \left(\frac{3}{2} + \frac{(-1)^n}{2}\right)\lambda^n + (1+n^2)e^{-n}. \tag{2.38}$$

The sequence β_n is bounded and has no limit, since it takes alternatively the values 2 and 1/2. Moreover, $\gamma = -1$. We consider three different cases:

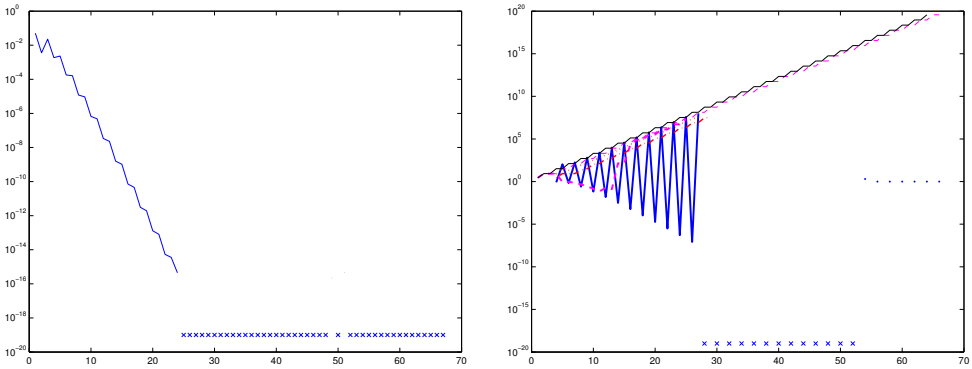
- $\lambda = \frac{1}{2}$ (Figure 2.8a): S_n is convergent, and $\lambda = \liminf_{n\to\infty} \beta_n$ (hence an accumulation point);
- $\lambda = 2$ (Figure 2.8b): S_n is divergent, and $\lambda = \limsup_{n\to\infty} \beta_n$ (hence an accumulation point);
- $\lambda = \frac{9}{10}$ (Figure 2.8c): S_n is convergent, and $|\lambda - \beta_n| > \frac{1}{2}$ for any n (hence ${}^2\mathrm{T}_n \to S$ by Theorem 2.6).

The lack of a limit for β_n seems to not influence the convergence of the transformation. Indeed, we obtain good results in all cases. This means that the second condition of Theorem 2.5 is not a necessary condition.

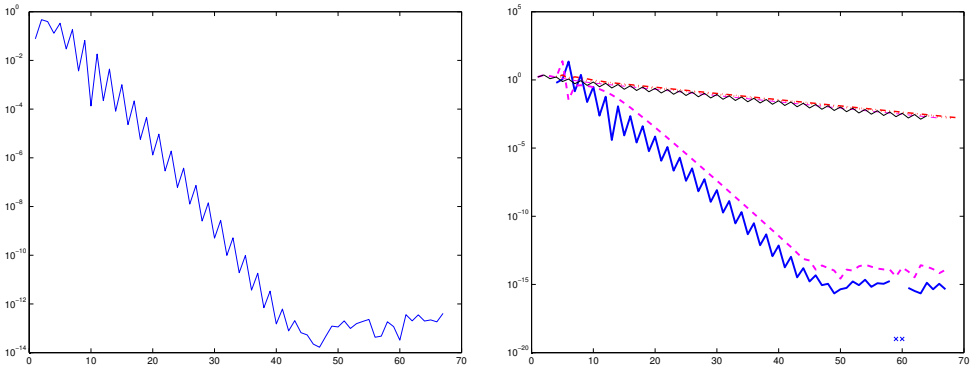
We underline that for $\lambda = 1/2$, the determinant of the system (2.4) is equal to $-3(S_{n+2} - S_n) - 2(S_{n+1} - S_n)$. Hence, when ΔS_n is close to the machine precision, we have singularity problems; see the circle in Figure



(a) $\lambda = 0.5$



(b) $\lambda = 2$



(c) $\lambda = 0.9$

Figure 2.8: On the left, the errors in the estimate of λ (obtained by system (2.4)) for sequence (2.38) with different values of λ. On the right, the values $|S - S_n|$ (solid) are compared with the errors obtained using Aitken's Δ^2 process (dashed), $\varepsilon_4^{(n)}$ (dashed bold), $\theta_2^{(n)}$ (dash-dotted), Levin type transformation (dotted) and transformation ${}^2\mathrm{T}_n$ (bold).

2.8a. However, this is not a problem since it happens when S_n reaches the value of S at machine precision.

Moreover, in Figure 2.8b when $\lambda_n = 2$ at machine precision, the denominator in transformation ${}^2\mathrm{T}_n$ is alternatively equal to 0 or 1. Hence, after the first iterations ${}^2\mathrm{T}_n$ is not computed for n odd.

Finally, all the other transformations considered do not accelerate any of these sequences, except for $\varepsilon_4^{(n)}$ which accelerates the first case, semi-converges in the last case, and partially in the second one. However, we remark that ${}^2\mathrm{T}_n$ uses less terms of the sequence than $\varepsilon_4^{(n)}$.

Summarizing, in this chapter we introduced three new transformations which accelerate the convergence of sequences that are not far from the ones of the form of (2.1), with a_n and S_n given sequences. Numerical results showed that transformation ${}^2\mathrm{T}_n$ performed better than the other two on the example we considered. Moreover, we compared ${}^2\mathrm{T}_n$ with several well-known transformations, obtaining competitive results under the assumptions of the convergence and acceleration theorems proved in Section 2.2. However, numerical experiments showed that these conditions are not necessary.

In addition, we compared ${}^2\mathrm{T}_n$ with the transformations introduced in [13]. We get some good results. Nevertheless, it is important to remark that ${}^2\mathrm{T}_n$ consider more information about sequence (S_n) than the other ones.

2.4 Accelerating Gauss quadrature, some prospectives

We present some ideas on possible new developments. Let us consider the kernel defined by the following kind of sequences

$$S_n = S + a_n(\lambda_n)^n, \tag{2.39}$$

with S unknown, a_n a known sequence and λ_n an unknown sequence. Clearly it extends kernel (2.1) in which λ_n is a constant sequence. It could be of interest to give a transformation similar to ${}^2\mathrm{T}$ in order to accelerate sequences closed to kernel 2.39.

In Part I Chapter 4 we deeply discuss Gauss quadrature formula for the approximation of integrals and linear functionals. In particular, let f be

a continuous function and

$$I(f) = \int_{\mathbb{R}} f \mathrm{d}\mu,$$

the Riemann-Stieltjes integral with respect to a non-decreasing distribution function μ defined on the real axis having finite limits at $\pm\infty$ and infinitely many points of increase. Let $G_n(f)$ be the n-node Gauss quadrature approximating $I(f)$ (see (4.2) in Part I), and $\theta_1 < \cdots < \theta_n$ be its nodes. Then as shown for example in [20] we can express its error as

$$E_n(f) = I(f) - G_n(f) = \frac{f^{(2n)}(\xi_n)}{2n!} I(\pi_n^2),$$

with $\theta_1 < \xi_n < \theta_n$ and π_n the $n+1$ monic orthogonal polynomial with respect to I (for a definition see Part I Chapter 1). If we consider $f(x) = x^k$ with k an integer, then

$$E_n(f) = \left(\begin{array}{c} k \\ 2n \end{array} \right) I(\pi_n^2) \left(\frac{1}{\xi_n} \right)^{2n-k}.$$

Hence, we can rewrite G_n as

$$G_n(x^k) = I(x^k) - \left(\begin{array}{c} k \\ 2n \end{array} \right) I(\pi_n^2) \left(\frac{1}{\xi_n} \right)^{2n-k}.$$

The Gauss quadrature G_n so expressed is of the type of (2.39), with

$$\begin{aligned} S_n &= G_n(x^k), \\ S &= I(x^k), \\ a_n &= -\left(\begin{array}{c} k \\ 2n \end{array} \right) I(\pi_n^2), \\ \lambda_n &= \frac{1}{\xi_n^{2-k/n}}. \end{aligned}$$

Therefore, a transformation similar to ${}^2\mathrm{T}_n$ could be useful for the acceleration of the sequence of Gauss quadrature $G_1, G_2, \ldots$. However, the behavior of λ_n is not clear and its estimation could be more difficult then in the case we have considered in the previous chapters. Once we find a way to accelerate $G_n(x^k)$ we can try to extend it to any $G_n(f)$ with f a polynomial. Finally, we may consider an analogous study for the case of n-weight Gauss quadrature for the approximation of quasi-definite linear functionals (see Part I Section 4.2).

Bibliography

[1] J. Abdalkhani, D. Levin, On the choice of β in the u - transformation for convergence acceleration, submitted.

[2] M. Abramowitz, I.A. Stegun, Handbook of mathematical functions, National Bureau of Standards, Washington, D. C., 1972.

[3] A.C. Aitken, On Bernoulli's numerical solution of algebraic equations, Proc. R. Soc. Edinb. 46 (1926) 289–305.

[4] B. Beckermann, A connection between the E-algorithm and the epsilon algorithm, In: C. Brezinski (Ed.), Numerical and Applied Mathematics, vol. 1.2, J. C. Baltzer, Basel, 1989, pp. 443–446.

[5] F. Bornemann, D. Laurie, S. Wagon, J. Waldvogel, The SIAM 100-Digit challenge: A study in high-accuracy numerical computing, Society for Industrial and Applied Mathematics, Philadelphia (2004).

[6] C. Brezinski, Accélération de suites à convergence logarithmique, C. R. Acad. Sci. Paris 273 A (1971) 727–730.

[7] C. Brezinski, Padé-type approximation and general orthogonal polynomials, Birkhäuser-Verlag, Basel, 1980.

[8] C. Brezinski, The asymptotic behavior of sequences and new series transformations based on the Cauchy product, Rocky Mountain J. Math. 21 (1991) 71–84.

[9] C. Brezinski, Convergence acceleration during the 20th century, J. Comput. Appl. Math, 122 (2000) 1–21.

[10] C. Brezinski, M. Crouzeix, Remarques sur le procédé Δ^2 d'Aitken, C. R. Acad. Sci. Paris 270 A (1970) 896–898.

[11] C. Brezinski, M. Redivo-Zaglia, Construction of extrapolation processes, Appl. Numer. Math. 8 (1991) 11–23.

[12] C. Brezinski, M. Redivo-Zaglia, Extrapolation methods: Theory and practice, North-Holland, Amsterdam, 1991.

[13] C. Brezinski, M. Redivo-Zaglia, Generalizations of Aitken's process for accelerating the convergence of sequences, Comput. Appl. Math. 26 (2007) 171–189.

[14] D. Buoso, A. Karapiperi, S. Pozza, Generalizations of Aitken's process for a certain class of sequences, Appl. Numer. Math. 90 (2015) 38–54.

[15] J.P. Delahaye, Sequence transformations, Springer-Verlag, Berlin, 1988.

[16] J.P. Delahaye, B. Germain-Bonne, Résultats négatifs en accélération de la convergence., Numerische Mathematik 35 (1980) 443–457.

[17] J.P. Delahaye, B. Germain-Bonne, The set of logarithmically convergent sequences cannot be accelerated, SIAM J. Numer. Anal. 19 (1982) 840–844.

[18] T. Fessler, W.F. Ford, D.A. Smith, HURRY: An acceleration algorithm for scalar sequences and series, ACM Trans. Math. Software 9 (1983) 346–354

[19] A. Gil, J. Segura, N.M. Temme, Numerical methods for special functions, SIAM, Philadelphia, 2007.

[20] G.H. Golub, G.A. Meurant, Matrices, moments, and quadrature with applications, Princeton University Press, Princeton, N.J., 2010.

[21] C. Kowalewski, Accélération de la convergence pour certaines suites à convergence lagarithmique, In: M.G. de Bruin, H. van Rossum (Ed.), Padé Approximation and its applications, vol. 888 LNM, Springer-Verlag, Berlin, 1981, pp. 263–272.

[22] C. Kowalewski, Possibilités d'accélération de la convergence logarithmique, Thèse 3ème cycle, Université de Lille I, 1981.

[23] D. Levin, Development of non-linear transformations for improving convergence of sequences, Int. J. Comput. Math. B 3 (1973), 371–388.

[24] F.W.J. Olver, D.W. Lozier, R.F. Boisvert, C.W. Clark, NIST handbook of mathematical functions, Cambridge U. P., Cambridge, 2010.

[25] N. Osada, The early history of convergence acceleration methods, Numer. Algor. 60 (2012) 205–221.

[26] W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, Numerical recipes: The art of scientific computing, Cambridge University Press, Cambridge, 2007.

[27] D. Shanks, Non linear transformations of divergent and slowly convergent sequences, J. Math. Phys. 34 (1955) 1–42.

[28] A. Sidi, Practical extrapolation methods, Cambridge University Press, Cambridge, 2003.

[29] T.J. Stieltjes, Recherches sur quelques séries semi-convergentes, Ann. Sci. Ec. Norm. Super. 3 (1886) 201–258.

[30] L.N. Trefethen, Approximation theory and approximation practice, SIAM, Philadelphia (2013).

[31] R.R. Tucker, The δ^2-process and related topics, Pacific J. Math. 22 (1967) 349–359.

[32] E.J. Weniger, Nonlinear sequence transformations for the acceleration of convergence and the summation of divergent series, Comput. Physics Reports 10 (1989) 189–371.

[33] E.J. Weniger, Irregular input data in convergence acceleration and summation processes: General considerations and some special Gaussian hypergeometric series as model problems, Comput. Phys. Commun. 133 (2001) 202–228.

[34] E.J. Weniger, A rational approximant for the digamma function, Numer. Algor. 33 (2003), 499–507.

[35] E.J. Weniger, Further discussion of sequence transformation methods, Subtopic "Related Resources" (R1) on the Numerical Recipes (Third Edition) Webnotes page http://www.nr.com/webnotes/ (2007).

[36] E.J. Weniger, On the mathematical nature of Guseinov's rearranged one-range addition theorems for Slater-type functions, J. Math. Chem. 50 (2012) 17–81.

[37] J. Wimp, Sequence transformations and their applications, Academic Press, New York, 1981.

[38] P. Wynn, On a device for computing the $e_m(S_n)$ transformation, Math. Tables Aids Comput. 10 (1956), 91–96.